IJES International Journal of English Studies

Cognitive Phonology

Issue Editor:

José Antonio Mompeán

Volume 6 Number 2 2006

INTERNATIONAL JOURNAL OF ENGLISH STUDIES

Volume 6, Number 2, 2006

Monograph:

Cognitive Phonology

Issue Editor:

José Antonio Mompeán

General Editor

Pascual Cantos Gómez (University of Murcia)

Issue Editors

José Antonio Mompeán (University of Murcia)

Editorial Assistants

Keith Gregor (University of Murcia), *Juan Antonio Suárez* (University of Murcia)
David Walton (University of Murcia), *Clara Calvo* (University of Murcia)
Rosa M. Manchón (University of Murcia), *Javier Valenzuela* (University of Murcia)

Editorial Advisory Board

Eva Alcón (Jaume I University, Castellón, Spain)
Doug Arnold (University of Essex, England)
David Britain (University of Essex, England)
Robert Carlisle (California State University, USA)
Jack K. Chambers (University of Toronto, Canada)
Rebecca Clift (University of Essex, England)
Ángeles de la Concha (UNED, Spain)
René Dirven (University of Duisburg, Germany)
Fred Eckman (University of Wisconsin–Milwaukee, USA)
Alvino Fantini (International School of Intercultural Training, Vermont, USA)
Francisco Fernández (University of Valencia, Spain)
Fernando Galván Reula (University of Alcalá, Spain)
Francisco Garrudo (University of Seville, Spain)
Rosa González (University of Barcelona, Spain)
Santiago González Fernández-Corugedo (University of Oviedo, Spain)
Michael Hattaway (University of Sheffield, England)
Leo Hickey (University of Salford, England)
Ton Hoenselaars (University of Utrecht, Holland)
Richard Hogg (University of Manchester, England)
Glenn Jordan (University of Glamorgan, Wales)
Vijay Kumar Bhatia (Hong Kong City University, China)
George Lakoff (University of California at Berkeley, USA)
Roger Lass (University of Capetown, South Africa)
Terttu Nevalainen (University of Helsinki, Finland)
Susana Onega (University of Saragossa, Spain)
Gary Palmer (University of Nevada at Las Vegas, USA)
Klaus-Uwe Panther (University of Hamburg, Germany)
Lionel Pilkington (NUI Galway)
Rafael Portillo (University of Seville, Spain)
Günter Radden (Hamburg University, Germany)
Helena Raumolin-Brunberg (University of Helsinki, Finland)
A. Robert Lee (Nihon University, Japan)
Françoise Salager-Meyer (Los Andes University, Venezuela)
Alexander Shurbanov (University of Sofia, Bulgaria)
John Storey (University of Sunderland, England)
Paul Tench (University of Wales Cardiff)
Linda Thornburg (University of Hamburg, Germany)
Mª Teresa Turell (Pompeu Fabra University, Spain)
Peter Trudgill (University of Fribourg, Switzerland)
Andrew Varney (University of Wales Swansea)
Chrish Weedon (University of Wales Cardiff)

Fernando Justicia (University of Granada, Spain)

The *International Journal of English Studies* and its contents are indexed in MLA (Modern Language Association) Directory of Periodicals, the ERIC (Educational Resources Information Database) Database and the IBSS (International Bibliography of the Social Sciences).

Editorial policy

The *International Journal of English Studies* (IJES), published by the University of Murcia (Spain), is a refereed journal which seeks to reflect developments in the general field of English Studies. Edited by members of Murcia's Department of English, the Journal is published twice-yearly in the form of English-language monographs in the fields of Language and Linguistics, Language Learning and Teaching, Literature and Cultural Studies, and attracts contributions from both nationally and internationally acclaimed scholars.

Subscriptions, Advertising and Exchange

The subscription rates and single issues are as follows (including postage and VAT):

One-year subscription:	15 € (15 $) individuals;	30 € (30 $) institutions
Single issue:	10 € (10 $) individuals;	20 € (20 $) institutions

We regret that we cannot accept credit card payments yet. Please send the **Order Form** that you will find at http://www.um.es/engphil/ijes, together with a cheque payable to **UNIVERSIDAD DE MURCIA**, either in Euros or in US Dollars, to:

Cristina Soriano (csoriano@um.es) or Flor Mena (flormena@um.es)
Departamento de Filología Inglesa
Facultad de Letras
Universidad de Murcia
Campus de La Merced
30071 Murcia (Spain)

Copyright

All rights reserved. No part of this publication may be reproduced, in any form or by any means (photocoping, electronic, or otherwise) without written permission from the Servicio de Publicaciones of the University of Murcia. In assigning copyright the author retains his or her personal right to re-use material in future collections of his or her own work with no fee to the Journal. Proper acknowledgement of prior publication in *International Journal of English Studies* is the only requirement in such cases.

Notice

No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. All contributions are anonymously refereed by three specialists in the field(s).

© Servicio de Publicaciones. Universidad de Murcia.

ISSN: 1578-7044
D.L.: MU-2014-2001

Table of Contents

IJES, Volume 6, Number 2, 2006

INTRODUCTION: *Cognitive Phonology in Cognitive Linguistics*..... vii
JOSÉ A. MOMPEÁN

ARTICLES:

DAVID EDDINGTON
‹‹Paradigm Uniformity and Analogy: The *Capitalistic* versus *Militaristic* Debate››……... 1

JOHN R. TAYLOR
‹‹Where do Phonemes Come from? A View from the Bottom››…….... 19

HELEN FRASER
‹‹Phonological Concepts and Concept Formation: Metatheory, Theory and Application››……. 55

FUMIKO KUMASHIRO & TOSHIYUKI KUMASHIRO
‹‹Interlexical Relations in English Stress››……..... 77

GITTE KRISTIANSEN
‹‹Towards a Usage-Based Cognitive Phonology››…..…….... 107

JOSÉ A. MOMPEÁN
‹‹The Phoneme as a Basic-Level Category: Experimental Evidence from English››……..... 141

GEOFFREY S. NATHAN
‹‹Is the Phoneme Usage-Based? – Some Issues››…….……..... 173

BOOK REVIEWS:

‹‹Review of *Cognitive Phonology in Construction Grammar: Analytic Tools for Students of English* by RiittaVälimaa-Blum (2005). Berlin: Mouton de Gruyter››……..... 195

ABOUT THE AUTHORS…..……..... 201

INSTRUCTIONS TO AUTHORS……... 205

ALREADY PUBLISHED AND FORTHCOMING NUMBERS……... 211

International Journal of English Studies

IJES

UNIVERSITY OF MURCIA www.um.es/engphil/ijes

Cognitive Phonology in Cognitive Linguistics

As it has generally been conceived of since its inception, cognitive linguistics (CL) is an approach to the study of language that endeavours to explain facts about language in terms of known properties and mechanisms of the human mind/brain. The guiding principle behind this area of linguistics is that the human language ability is not separate from the rest of cognition, that the storage and retrieval of linguistic data is not significantly different from the storage and retrieval of other knowledge, and that use of language in understanding employs similar cognitive abilities to those used in other non-linguistic tasks. CL also argues that language is embodied and situated in the sense that it is embedded in the experiences and environments of its users.

Since CL was 'officially' born in the mid-eighties with the seminal works by Lakoff (1987) and Langacker (1987), most studies within this school of linguistics have focused on semantics and grammar. It is true that even the work by Lakoff and Langacker already contains references to phonology (in fact, Langacker's Cognitive Grammar gives a prominent role to the phonological pole of linguistic units). However, phonological work in CL, with notable exceptions such as the work of a few scholars like Geoffrey S. Nathan (e.g. Nathan, 1986, 1994, 1996, 1999) or John Taylor (e.g. Taylor, 1989, 1990, 2002), has always been sparse in comparison with the attention paid to other areas of study like semantics or grammar. This is even more surprising since the study of phonology represented, historically, the onset of modern linguistics and was also a flagship of other approaches to language like structuralist (e.g. Trubetzkoy, 1939) or generative linguistics (e.g. Chomsky & Halle, 1968).

Despite the traditional under-representation of phonological issues within CL, recent years have witnessed a growing interest in phonology within CL, shown by the increasing number of phonological papers in books or journals like *Cognitive Linguistics* and the phonology theme sessions organized at the *International Cognitive Linguistics Association* conferences from 1999

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

on. One important aspect of this increasing body of work is that it is heterogeneous as to the formalisms and methods that it uses, which does not make it a rigid field (from the formalist point of view) like other recent approaches to phonology (e.g. Optimality Theory) or one defined by reference to the methods that it uses (e.g. Laboratory or Experimental Phonology).

Despite the heterogenity of current cognitive phonological work, researchers working in this area endorse the view that the phonological component of languages can be explained with reference to known properties and mechanisms of the human mind/brain as well as by embodied experience and environmental factors. 'Cognitive phonologists' also endorse the *Cognitive Commitment* and the *Generalization Commitment* (see e.g. Gibbs, 1996; Lakoff, 1990). Applied to phonology, the *Cognitive Commitment* implies that phonological concepts, categories and constructs need to have psychological validity, which is best secured by informing phonological work with a broad empirical basis -including experimental research- from a wide range of disciplines like phonetics, psycholinguistics, sociolinguistics, second language acquisition, cognitive psychology, developmental psychology, etc. The *Generalization Commitment* implies that, although it may be useful to treat different areas of language study (e.g. phonology, semantics, syntax, morphology, etc.) as notionally distinct, CL is committed to investigating "how the various aspects of linguistic knowledge emerge from a common set of human cognitive abilities upon which they draw, rather than assuming that they are produced in encapsulated modules of the mind" (Evans *et al.*, in press: 5). In addition, 'cognitive' phonologists also endorse the view that phonological categories/constructs are shaped not only by cognitive factors but also by factors of a phonetic (articulatory, acoustic, perceptual), linguistic (historical, distributional, structural, frequential, etc.), sociolinguistic (gender, social class, etc.), cultural (e.g. orthographic), or developmental kind (factors of other kinds can also be added to this list).

It is within this context that this special volume of IJES makes it appearance. By bringing together a number of papers on the segmental and suprasegmental aspects of language (English as the reference language) and especially written for this occasion, the volume aims to contribute to the relative scarcity of work on phonology within CL.

David Eddington opens the volume with his paper "Paradigm Uniformity and Analogy: The *Capitalistic* versus *Militaristic* Debate", in which he presents the results of a study in which the author used a computationally explicit algorithm and a data set of 3,719 instances of the allophones of /t/ taken from a language corpus. The results obtained lead to the author's

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

contention that all allophonic distribution may be explained in terms of analogy to stored linguistic experience. This is in contrast to previous ideas of analogy as a process that interferes with the application of general rules.

The contribution by **John Taylor** entitled "Where do Phonemes Come from? A View from the Bottom" explores the possible sources of phonemes as conceptual categories. In this paper, the author contrasts two learning paradigms -supervised learning (where learners receive feedback on their categorization attempts) and unsupervised learning (where learners rely only on properties of the input). Taylor argues that unsupervised learning may be the appropriate paradigm, at least for the initial stages of acquisition of phonological categories. Thereafter, the emergence of phoneme categories draws on various kinds of knowledge available to the learner, including knowledge of articulation, and of literacy conventions. The paper concludes with a section that emphasizes the taxonomic nature of the phoneme, and suggests that the special salience of a phonemic representation reflects the status of the phoneme as a basic-level category.

The paper by **Helen Fraser**, "Phonological Concepts and Concept Formation: Metatheory, Theory and Application" presents an overview of Phenomenological Phonology, (including its metatheory, theory and application) for comparison with Cognitive Phonology. The author claims that, while Phenomenological Phonology and Cognitive Phonology are in close agreement at the theory level, there are some significant differences at the level of metatheory. As a case in point, Phenomenological Phonology considers phonological terms (such as *phoneme* and *word*) to be words like any others, and gives detailed consideration to the concepts behind such terms. It also considers pronunciation to be a form of behaviour, driven by concepts created through general concept-formation processes. This has important consequences for practical application in the areas of pronunciation and literacy teaching.

Within Ronald Langacker's theory of Cognitive Grammar, **Fumiko Kumashiro** and **Toshiyuki Kumashiro**, in their article "Interlexical Relations in English Stress", propose a cognitive, non-reductionist analysis of English stress as it pertains to interlexical relations, based on the usage-based model as proposed by Cognitive Grammar and on the connectionist interactive activation model. The authors claim that interlexical relations involved in English stress can be satisfactorily accounted for by employing actually-occurring expressions as

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

constraints and that precise explication of these relations requires consideration not only of phonological but also of semantic factors.

Gitte Kristiansen's paper "Towards a Usage-Based Cognitive Phonology" argues that cognitive phonology must aim at a higher degree of descriptive refinement, especially in the direction of social variation. The paper argues that phonemic analysis should be carried out taking into account the rich patterns of language-internal variation. It also examines the implications of a usage-based and multi-faceted model for a theoretical discussion of the phoneme as a prototype category.

In the paper entitled "The Phoneme as a Basic-Level Category: Experimental Evidence from English", **José A. Mompeán** presents the results of a concept formation experiment that provides evidence on the possible existence of a basic-level of taxonomic organization in phonological categories as conceived of by phonetically naïve, native speakers of English. This level is roughly equivalent to the phoneme as described by phonologists and linguists. The paper also discusses the reasons why the phoneme could be considered as the basic level of taxonomies of phonological categories.

The paper "Is the Phoneme Usage-Based: Some Issues" by **Geoffrey S. Nathan** presents a brief review of the history of the phoneme, from its origins in the nineteenth century to Optimality Theory, including some Cognitive Linguists' views of the concept. In the paper, Nathan argues that current 'usage-based' theorists views of the phoneme may not be able to explain some facts about how naïve speakers process language, both consciously and subconsciously. These facts include the invention of and worldwide preference for alphabetic writing systems, and language processing evidence provided by Spoonerisms, historical sound changes affecting all (or most) lexical items in a language and each other, and the fact that allophonic processes normally do not show lexical conditioning. The author further suggests that storing speech in terms of a small number of production/perception units such as phonemes could be due to the fact that phonemes seem to optimize both efficiency and informativeness in much the same way as other basic-level categories.

A review of a recently published book brings this monograph issue of IJES to a close. **John R. Taylor** reviews Riita Välimaa-Blum's recent book *Cognitive Phonology in Construction Grammar: Analytic Tools for Students of English.*

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I would like to end this introduction by expressing my gratitude to all the contributors to the volume for their professionalism and patience in the process of editing this monograph, as well as to the referees who evaluated the texts and supplied valuable feedback and advice to the authors. I expect that readers will find this collection an interesting sample of how cognitive approaches to language can be applied to phonological work and that this collection will contribute both to the interest in the study of phonology among cognitive linguists and the interest among phonologists in cognitive approaches to phonological work.

JOSÉ A. MOMPEÁN
Issue Editor

REFERENCES

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Evans, V., Bergen, B. K., & Zinken, J. (in press). The Cognitive Linguistics enterprise: An overview. In *The Cognitive Linguistics Reader*. London: Equinox.

Gibbs, R. W. (1996). What's cognitive about cognitive linguistics?. In E. H. Casad (Ed), *Cognitive linguistics in the redwoods.* Berlin/New York: Mouton de Gruyter, pp. 27-53.

Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, G. (1990). The invariance hypothesis: Is abstract reason based on image-schemas?. *Cognitive Linguistics*, 1(1), 39-74.

Langacker, R. (1987). *Foundations of cognitive grammar, volume 1, Theoretical prerequisites.* Stanford: Stanford University Press.

Nathan, G. S. (1986). Phonemes as mental categories. *Berkeley Linguistics Society,* 12, 212-223.

Nathan, G. S. (1994). How the phoneme inventory gets its shape: Cognitive grammar's view of phonological systems. *Rivista di Linguistica,* 6, 275-287.

Nathan, G. S. (1996). Steps towards a cognitive phonology. In B. Hurch & R. Rhodes (Eds.), *Natural phonology: The state of the art*. Berlin: Mouton de Gruyter, pp. 107-120.

Nathan, G. S. (1999). What functionalists can learn from formalists in phonology. In M. Darnell *et al.* (Eds.), *Functionalism and formalism in linguistics: Volume I: General papers*. Amsterdam/Philadelphia: John Benjamins, pp. 305-327.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Taylor, J. R. (1989) *Linguistic categorisation: Prototypes in linguistic theory*. Oxford. Oxford University Press. 2nd edition, 1995, 3rd edition 2003.

Taylor, J. R. (1990). Schemas, prototypes, and models: In search of the unity of the sign. In S. L. Tsohatzidis (Ed.). *Meanings and prototypes. Studies in linguistic categorisation*. London: Routledge, pp. 521-534.

Taylor, J. R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.

Trubetzkoy, N. (1939). *Grundzüge der Phonologie*. Travaux du Cercle Linguistique de Prague 7.

Välimaa-Blum, R. (2005). *Cognitive phonology in construction grammar: Analytic tools for students of English*. Berlin: Mouton de Gruyter

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Paradigm Uniformity and Analogy: The Capitalistic versus Militaristic Debate[1]

DAVID EDDINGTON*
Brigham Young University

ABSTRACT

In American English, /t/ in *capi*t*alistic* is generally flapped while in *mili*t*aristic* it is not due to the influence of *capi[ɾ]al* and *mili[*tʰ*]ary*. This is called Paradigm Uniformity or PU (Steriade, 2000). Riehl (2003) presents evidence to refute PU which when reanalyzed supports PU.

PU is thought to work in tandem with a rule of allophonic distribution, the nature of which is debated. An approach is suggested that eliminates the need for the rule versus PU dichotomy; allophonic distribution is carried out by analogy to stored items in the mental lexicon. Therefore, the influence of the pronunciation of *capital* on *capitalistic* is determined in the same way as the pronunciation of /t/ in monomorphemic words such as *Mediterranean* is. A number of analogical computer simulations provide evidence to support this notion.

KEYWORDS: analogy, flapping, tapping, American English, phoneme /t/, Analogical Modeling of Language, allophonic distribution, Paradigm Uniformity.

* ***Address for correspondence***: David Eddington. Brigham Young University. Department of Linguistics and English Language. 4064 JFSB. Provo, UT 84602. Phone: (801) 422-7452. Fax: (801) 422-0906. e-mail: eddington@byu.edu

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I. INTRODUCTION

In traditional approaches to phonology, all surface forms are generated from abstract underlying forms. However, the pre-generative idea that surface forms can influence other surface forms, (e.g. paradigmatic analogy in historical linguistics) has reemerged in a number of formal models (Benua, 1995; Burzio, 1996; Kenstowicz, 1996; McCarthy, 1995; McCarthy & Prince 1994a, b; Steriade, 1997, 1999, 2000). The differences in the American English pronunciation of the medial /t/ in *capitalistic* and *militaristic* (*mili[*tʰ*]aristic* versus *capi[ɾ]alistic*) has received a great deal of attention in this regard beginning with Withgott (1982). Since /t/ appears in the same phonetic environment in both words, it should be given the same pronunciation. The fact that it is generally flapped[2] in *capitalistic* but aspirated in *militaristic* is thought to be due to the pronunciation of the base words *mili[*tʰ*]ary* and *capi[ɾ]al.*

Steriade (2000) accounts for these words by appealing to the notion of Paradigm Uniformity. For Steriade, there are two competing processes; regular phonological distribution is the default that is occasionally interrupted by the effects of Paradigm Uniformity (henceforth PU). In the present paper, two criticisms of her analysis of *capitalistic* and *militaristic* are discussed. The discussion is couched in terms of an explicit model of linguistic analogy that holds that allophonic distribution is based on analogy to stored memory tokens of past linguistic experience. In contrast to Steriade's notion of regular distribution plus PU, analogy is a unitary process that predicts all instances of phonological distribution, not just the exceptional cases that appear to be due to the influence of other members of a paradigm.

II. PARADIGM UNIFORMITY

According to Paradigm Uniformity (Steriade, 1997, 1999, 2000) if a base has a particular non-contrastive phonetic feature, derivatives of that base will tend to keep that feature. Since *military* contains a medial [tʰ] while the /t/ in *capital* is generally flapped, these features carry over into the derivatives *mili[*tʰ*]aristic* and *capi[ɾ]alistic* in spite of the fact that /t/ appears in a similar phonetic context in both of the derived words. Steriade (2000) tested the idea of PU by having

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

subjects read a list of ten words, some of which are generally flapped (e.g. *rotary*) and others that are not (e.g. *voluntary*), and then having them read a list of neologisms ending in *-istic* based on the ten words (e.g. *voluntaristic*, *rotaristic*). She found that 11 of the 12 subjects pronounced the derived forms with the same phone (i.e. either [tʰ] or [ɾ]) as they did the base forms, which she presents as evidence in favor of PU.

However, Riehl (2003) takes issue with Steriade's neologism study. She replicated Steriade's experiment with the modification that the test subjects repeated each of the four base and four derived forms (*negative/istic, positive/istic, primitive/istic, relative/istic*) twelve times rather than once. According to Riehl, PU would only be supported if all 12 repetitions of a base and its derived form are produced with [ɾ] or if none of the 12 are pronounced with a flap. In her study, some variability was encountered within a single subject's responses (e.g. *primi[*tʰ*]ivistic* vs. *primi[ɾ]ivistic*) which Riehl takes as a prima facie refutation of PU.

Riehl is correct in pointing out that Steriade does not clarify how variation in pronunciation would fit into PU. However, Riehl's data may actually support PU if statistical tendencies, rather than an all or nothing interpretation, are considered. To this end, a correlation was performed using Riehl's experimental results; the number of times each speaker used a flap in the base form was correlated with the number of times a flap was used in the derived form. The analysis comprised the data for all four test items and was highly significant (r (14) = .748, $p < .0005$, two-tailed).[3] This demonstrates that the more often a speaker flapped the /t/ in the base form, the more often he or she flapped the /t/ in the derived word, and vice-versa. However, the one-to-one correspondence that Riehl seemed to expect was not present. Riehl's findings are easily accounted for in terms of analogy, but an introduction to the particular model of analogy espoused in the present paper is in order before proceeding any further.

III. ANALOGY

In traditional approaches, analogy has been used to patch up the cases that the operation of supposedly regular processes fails to account for. According to the model employed in the

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

present paper (i.e. Analogical Modeling; Skousen, 1989, 1992, 1995, 1998), linguistic processing, including phonological distribution, is the result of analogy. The distribution of the allophones of /t/ is generally thought to be a matter of finding which context each allophone occurs in, storing the generalizations gleaned from the input, and then applying them in subsequent linguistic processing. Formal accounts differ most in terms of what factors and mechanisms they allow in deriving the correct allophone in the correct context (Davis, 2005; Giegerich, 1992; Harris, 1994; Jensen, 1993; Kahn, 1980; Kiparsky, 1979; Nespor & Vogel, 1986; Rhodes, 1994; Selkirk, 1982). Analogy, on the other hand, assumes that speakers store their linguistic experience in all of its redundant glory, a notion that is supported by a great deal of empirical evidence (Alegre & Gordon, 1999; Baayen, Dijkstra, & Schreuder, 1997; Bod, 1998; Brown & McNeill, 1966; Bybee, 1994, 1995, 1998; Goldinger, 1997; Manelis & Tharp, 1977; Palmeri, Goldinger & Pisoni, 1993; Pawley & Syder, 1983; Pisoni, 1997; Sereno & Jongman, 1997). The idea that behavior is influenced by analogy to past experience has been demonstrated with both linguistic and non-linguistic data (e.g. Bybee & Slobin, 1982; Chandler, 1995, 2002; Hintzman, 1986, 1988; Hintzman & Ludlam, 1980; Stemberger & MacWhinney, 1988).

According to analogy, exactly which stored instances will influence the choice of allophone depends on similarity.[4] The choice of which phone to use in a word such as *capitalistic* is influenced by analogy from a number of different sources, the largest one being the word *capital* because it has so many orthographic, semantic, and phonetic characteristics in common with *capitalistic.* It is safe to assume that most instances of *capital* in an American English speaker's mental lexicon contain a flap but some also may contain [tʰ].[5] This, in and of itself, accounts for some of the variability registered by Riehl. Although *capital* is arguably the most prominent analog for *capitalistic,* any stored instance of a word that has characteristics in common with *capitalistic* can exert some influence. The more the two have in common, the greater the chances of influence. In all likelihood, some of the analogs would point to a [tʰ] pronunciation and others to a [ɾ] pronunciation.

If the sum influence from all relevant stored tokens pointed to 90% [ɾ] and 10% [tʰ] in *capitalistic* then there are at least two ways of using this stochastic knowledge (Skousen, 1989:

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

82). The first, called selection by plurality, is to consider the flap the "winner" and apply it in each case. The second, called random selection, is to consider the probabilities and apply them to the task at hand which would essentially result in pronouncing [ɾ] in 90% of the cases and [tʰ] in 10%. Using either random selection, selection by plurality, or both strategies, as children in non-linguistic experiments appear to do (Messick & Solley, 1957), would account for the sort of variability that occurs in actual language usage and that is hard to account for in formal approaches that predict one and only one outcome in a particular environment.

The question naturally arises regarding how closely this algorithm or any computer program models the mental mechanisms speakers employ in the course of language production. Analogy is based on the uncontroversial idea that linguistic information is stored in the mind and retrieved as necessary. That groups of similar words can effect the behavior of other words with similar characteristics is well-attested in the psycholinguistic literature (e.g. Bybee & Slobin, 1982; Stemberger & MacWhinney, 1988). There is also ample evidence that human behavior is based on stored exemplars (Eddington, 2000; Hall 2005; Medin & Schaffer, 1978; Murphy, 2002; Nosofsky, 1988: Schweitzer & Möbius, 2004; Solé, 2003). Computer algorithms of analogy are designed to model these effects. Therefore, what the brain and analogical models have in common is the ability to use a database of past experience to predict behavior. However, too little is known about the exact functioning of the brain to even begin to explain exactly how instances are stored, accessed, or categorized on the neural level. For this reason, it is impossible to conjecture about how faithfully any computer algorithm mirrors actual brain processes beyond the ability of both to analogize.

III.1. A simulation of derived *-istic* words

Rather than speculating about how analogy can account for the relationship between the phonetic shape of a base word and its *-istic* derivative, a concrete simulation was carried out in which the pronunciation of /t/ was predicted in a number of computer simulations.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

III.1.1.Test words

Test words included those used by Riehl and Steriade: *capitalistic, negativistic, positivistic, primitivistic,* and *relativistic.* All of these may be pronounced either with a flap or [tʰ] in American English. In addition, six other words with the same phonological structure have been included that are not part of the discussion in the extant literature on the topic: *habitability, irritability, immutability, dissatisfaction, concatenation*. *Concatenation* is interesting because rather than a flap, the base *concatenate* generally has a glottal stop ([kʰənˈkʰæʔnejt]).

III.1.2.The database

In order to carry out a simulation of these ten test words, a database is needed that represents past linguistic experience with /t/. A total of 3,719 instances of /t/ allophones were taken from the TIMIT corpus (Garofolo, Lamel, Fisher, Fiscus, Pallett, & Dahlgren, 1993; Zue & Seneff, 1996). TIMIT was originally designed for use in natural language processing tasks, and consists of 6,300 utterances resulting from having 630 speakers read 10 sentences each. There were 2,342 different sentences, some of which contained no instances of /t/ and others contained multiple instances. Past experience with analogical simulations shows that robust predictions are made on a database of a few thousand instances. For this reason, the TIMIT corpus was "mined" for the first 3,719 instances encountered. The phonetic transcription which was used in the simulations was carried out via acoustic analysis by the TIMIT researchers, however, they did not distinguish between released and aspirated allophones of /t/; instead, they indicated the voice onset time. For the purposes of the present paper, phones with a VOT of 60ms or higher were considered aspirated and those with a VOT of 59ms or lower as released unaspirated stops.

The resulting database used for the simulations contained 564 instances of [ɾ], 234 [ʔ], 284 [Ø], 760 [t], 860 [t⁻], and 969 [tʰ]. In addition, 48 instances of /t/ were voiced and much longer than a flap and were therefore transcribed as [d] in TIMIT (e.g. *carpenter, congratulate*).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

For each of the 3,719 instances of /t/, the particular allophone was identified. In instance (1) it appears as variable 1. The phonological and morphological context surrounding /t/ were converted into variables as well: the three phones or boundaries to the right of /t/ (variables 2-4) and the three phones or boundaries to the left of the /t/ (variable 5-7). The boundary values that could occupy one of the variable slots were either a phrase internal word boundary, a phrase internal pause, or a utterance initial or final pause/word boundary. The stress of the syllable preceding and following /t/ was also included (variables 8-9). For example, the flap pronunciation of /t/ in *meet* in the sentence *I know I didn't meet her…* yields this entry in the database:

(1) 1) ɾ, 2) word boundary, 3) m, 4) i, 5) word boundary, 6) ɚ 7) pause, 8) primary stress, 9) unstressed.

The simulation could have included semantic or orthographic variables as well; however, these variables proved sufficient for the purposes of the study.

III.1.3. Algorithm

The simulations were carried out using the algorithm in Analogical Modeling of Language (AM; Skousen, 1989, 1992, 1995, 1998). AM makes its predictions on the basis of a test item, which is a vector of variables that represents linguistic information about the entity whose behavior is being predicted (see Skousen, 1989, 1992 for a detailed treatment of the algorithm). A variable vector, such as in (1) above, contains information about the context in which /t/ occurs. If the goal is to predict the pronunciation of /t/ in (1), for example, the algorithm would search the database for items that share variables with (1), excluding of course the first variable which is the one that is being predicted. The algorithm then creates groups of database items with shared similarities called subcontexts. For example, one subcontext would contain all database items with [i] before /t/, another with [mi] preceding /t/, another with [i] preceding and a word

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

boundary following /t/, and so on until all single variables and combinations of variables are considered.

Variable vectors that have more in common with the test item will appear in more subcontexts. Subcontexts are further combined into more comprehensive groups called supracontexts. Some supracontexts will be homogenous in that members will "agree" and exhibit the same allophone of /t/ or the same variable vector. Other supracontexts will have disagreements in that they contain members with different allophones. Such supracontexts are heterogeneous and in some cases their members are eliminated from consideration as analogs (see Skousen, 1989). Minimizing disagreements by eliminating members of heterogenous supracontexts results in the analogical set, which can be conceived of as containing those database items that belong to the most clear-cut and unambiguous areas of contextual space.

AM uses the members of the analogical set to calculate the probability that the test item will be assigned one of the allophones of /t/ found in the database. Essentially, what AM calculates is that the allophone in the database items that are most similar to the test item will predict the behavior of the test item, although the allophones of /t/ that appear in less similar database items have a small chance of applying as well, provided that they appear in homogenous supracontexts. Allophony is always calculated in terms of a particular test item and as a result, no global characterization of the data is made as is the case for rules of allophony. This implies that the variables which may be important in determining the allophone of /t/ for one test item may be not be important in determining the allophone in a different one (Skousen, 1995: 223-226).

III.1.4. Method

In another study that used this algorithm and database (Eddington, 2007) analogy was able to correctly predict the pronunciation of /t/ in many of the database items. In addition, the rate of correct predictions remained high when analogs were drawn from only a small fraction of the database. Predictions also remained robust when supposedly critical variables such as stress were eliminated from consideration. In the present study, the model was used to predict the probability

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

that each of the allophones of /t/ would apply to the test words. Two sets of simulations were performed. Before carrying out the simulations, the base words (e.g. *capital, relative, immutable*) were deleted from the database along with other derivatives such as *capitalize* and *relatively*. In the flap simulations, one instance of the base word was then added to the database in which the pronunciation was a flap. In the aspiration simulations, the pronunciation was given as [tʰ] to each base word. For *concatenation*, one simulation was performed with [kʰənˈkʰæʔnejt] in the database and another with [kʰənˈkʰætʰənejt].

Test Word	**Simulation Type**	**tʰ**	**ɾ**	**t**	**Ø**	**ʔ**	**t¬**	**d**
capitalistic	Flapping Simulation	12	**78**	0	3	3	4	0
	Aspiration Simulation	**90**	0	0	3	3	4	0
negativistic	Flapping Simulation	1	**98**	0	0	0	1	0
	Aspiration Simulation	**93**	6	0	0	0	1	0
positivistic	Flapping Simulation	0	**99**	0	0	1	0	0
	Aspiration Simulation	**96**	3	0	0	1	0	0
primitivistic	Flapping Simulation	0	**96**	1	1	0	1	0
	Aspiration Simulation	**94**	2	1	1	0	1	0
relativistic	Flapping Simulation	10	**90**	0	0	0	0	0
	Aspiration Simulation	**86**	14	0	0	0	0	0
habitability	Flapping Simulation	8	**80**	0	0	0	14	0
	Aspiration Simulation	**80**	0	0	0	0	20	0
irritability	Flapping Simulation	3	**95**	0	0	0	2	0
	Aspiration Simulation	**96**	3	0	0	0	1	0
immutability	Flapping Simulation	4	**93**	0	1	0	3	0
	Aspiration Simulation	**83**	12	0	1	0	4	0
dissatisfaction	Flapping Simulation	0	**100**	0	0	0	0	0
	Aspiration Simulation	**100**	0	0	0	0	0	0
concatenation	Glottal Stop Simulation	0	0	0	0	**100**	0	0
	Aspiration Simulation	**100**	0	0	0	0	0	0

Table 1. Predicted probability of each allophone for simulations with [ɾ], [tʰ], or [ʔ] in the base form.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

III.1.5. Results of the simulations

The predicted probability of each allophone appears in Table 1. It should be clear that the simulations support the notion of PU; the pronunciation of the base word affects the pronunciation of the derived word. However, even when there is only one base form in the database with one outcome, and hence no variability, analogy predicts some slippage toward the other possible pronunciations. For example, when the base form is *nega[ɾ]ive*, *nega[ɾ]ivistic* is predicted at a rate of 98%, *nega[*t¬*]ivistic* at 1%, and *nega[*tʰ*]ivistic* at 1%. This was the sort of behavior that Riehl's subjects demonstrated. Another important point to mention is that the base word does not account for all of the analogical influence by itself. For example, when *capital* is included with the aspirate pronunciation, *capitalistic* is predicted to be aspirated at a rate of 90%. Inspection of the analogical set reveals that *capital* only accounts for 30% of that total. Eight other database items (e.g. *appetite, hepatitis,* and *particular)* account for the other 60% of the aspirated members of the analogical set. The remaining 10% is split between a deleted /t/ in *cent* and a glottal stop in one instance of *not*, and an unreleased pronunciation in the final /t/ of *participate*.

III.2. Simulation with morphologically simple words

According to Steriade (2000), the phonetic context in which /t/ appears in words such as *militaristic* and *capitalistic* favors a flap pronunciation, hence, the aspirate in *mili[*tʰ*]aristic* goes against the general trend due to PU. Davis (2005) argues that Steriade is wrong about the context of the flapping rule. In his analysis, aspirated stops reflect the general pattern. Therefore, *mili[*tʰ*]aristic* follows the regular distribution, and it is the flap in *capi[ɾ]alistic* that is unexpected and must be explained as due to the influence of the base *capi[ɾ]al*. To prove his point, Davis discusses a number of monomorphemic words that are phonologically similar to *militaristic* and *capitalistic* (e.g. *lollapalooza, abracadabra*). Two of these, *Mediterranean* and *Navratilova* contain medial /t/ and are directly relevant to the present discussion. Since these are pronounced

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

with aspirates rather than flaps, the general tendency for words of this sort must be [tʰ], and PU cannot play a role in their pronunciation because the words are monomorphemic.

III.2.1. Test words, database, algorithm, and method

Mediterranean and *Navratilova* were the test cases and the same database and algorithm described in sections III.1.2 through III.1.3 were used. No morphological relatives of the test words needed to be removed from the database prior to running the simulations, however.

Test Word	tʰ	ɾ	t	Ø	ʔ	t¬	d
Mediterranean	**68**	21	5	1	0	4	1
Navratilova	**47**	27	19	2	1	4	0

Table 2. Predicted probability of each allophone for monomophemic words.

III.2.2. Results of the simulations

The results appear in Table 2. Both words were predicted to have [tʰ] rather than [ɾ]. Does this mean that Davis' characterization is to be preferred over Steriade's since his default rules predict [tʰ]? Davis and Steriade both assume a framework in which generalizations about linguistic data are formulated and used in the course of language processing. Analogy to a paradigmatic relative is thought to override the generalization in certain cases. For Steriade, it overrides what should be a flap in *militaristic.* For Davis, it overrides what should be an aspirated stop in *capitalistic*. In other words, both researchers subscribe to the idea that analogy only plays a role in explaining exceptional cases not covered by the global generalization. In contrast to this view of analogy, the assumption underlying the present simulations is that no global generalizations about allophonic distribution are made, nor are they necessary. Instead, all predictions are made on a case-by-case basis. Analogy does not merely perform the task of accounting for exceptional outcomes due to paradigmatic similarity; it is used to predict all outcomes. Therefore, from an analogical perspective Davis is only correct as far as *Mediterranean* and *Navratilova* are concerned.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Whether his characterization is valid for other words with medial /t/, or for words containing medial stops other than /t/ would have to be determined separately.

IV. CONCLUSIONS

Riehl's study was designed to test whether PU could account for the discrepancy in the pronunciation of /t/ in words such as *capitalistic* and *militaristic*. Her test subjects did not behave in accordance with PU in 100% of the cases. However, a statistical analysis of those results reveals that the subjects' responses correlated highly with the predictions of PU, which actually argues in favor of Steriade's formulation of PU. A simulation of the words in question was performed using a computationally explicit model of analogy. The model predicts the sort of variability demonstrated by Reihl's subjects, and shows that analogical effects along the lines of PU are tenable but not void of variation.

Davis' critique of Steriade's analysis of *capitalistic* and *militaristic* concerns what allophone of /t/ should occur in the absence of PU. This would occur in monomorphemic words such as *Navratilova* and *Mediterranean*. Contra Steriade, he argues that [tʰ] is the default rather than [ɾ]. A simulation of the two monomorphemic words favors Davis' analysis; however, analogy works on a case-by-case basis and utilizes no global predictions, therefore, it can only verify Davis' analysis for these particular words. Both Davis and Steriade assume that an analogical process only applies when one surface form is a morphemic relative of another. In the rest of the instances, a more general process is thought to apply. The model of analogy described above, on the other hand, calculates all cases of allophony on the basis of stored memory traces. Accordingly, PU occurs because derived forms and their bases share many traits. Because analogy works on the basis of similarity, a base usually appears in the analogical set that is extracted from the mental lexicon when predicting the pronunciation of one of its derived forms. As a result, the base's pronunciation influences the pronunciation of the derived form.

There is a major advantage to the idea that allophonic distribution is carried out by analogy. Psychological evidence demonstrates that analogy plays an important role in human

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

cognition. In contrast, much of the machinery required in rule analyses has been called into question on formal grounds (Burzio, 1996; Cole, 1995; Cole & Hualde, 1998; Steriade, 1995). More importantly, the psychological reality of rules and constraints is highly questionable on empirical grounds as well (Derwing, 1973; Eddington, 1996; Lamb, 2000).

How could children subconsciously and effortlessly intuit the kinds of generalizations about allophony (which are often complex and abstract) that many intelligent graduate students of phonology have a difficult time formulating? If linguistic processing is analogical no such generalizations need to be made. If people formulate such generalizations, why are they not able to express them overtly? According to those who consider them psychologically real, it is because they learned and manipulated subconsciously. From an analogical viewpoint, speakers cannot describe the rule they use to determine that *plooty* would contain a flap because no rule exists. If pressed for an answer speakers will rarely give a rule-type response, but more often will state that *plooty* "sounds right" with a flap, or that it is similar to words such as *duty* and *booty*. Clearly, the search for psychologically plausible models of phonological processing must incorporate analogy.

NOTES

1. I express my thanks to Royal Skousen, José Antonio Mompeán, Dirk Elzinga, Andy Wedel, and Steve Chandler for their input on this paper.

2. According to very precise phonetic descriptions taps and flaps involve different articulations (Ladefoged, 2006: 170-171). In the present paper, the term flap is used to describe a non-retroflex, non-r-colored rapid stop gesture. This should cause no confusion since taps and flaps are not distinguished.

3. In some instances there is no data for a particular response. For example, one subject pronounced *positive* with a flap ten times and without a flap only once, and one response is missing. For the purposes of the correlation, 10.5 flap responses to *positive* were counted because the missing response would most likely have been another flap and that puts the figure halfway between the actual and probable number if all 12 responses had been given.

4. Exactly which characteristics are used to determine similarity is a question that needs to be explored in more depth. In this vein of research Eddington (2002) compared similarity based on phonemes versus similarity based on phonetic features and found no significant difference.

5. The fact that the word is written with a *t* surely plays a part as well beyond that of analogy.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

REFERENCES

Alegre, M. & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language,* 40, 41-61.

Baayen, R. H. Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language,* 37, 94-117.

Benua, L. (1995). Identity effects in morphological truncation. *Papers in Optimality Theory. University of Massachusetts Occasional Papers in Linguistics,* 18, 77-136.

Bod, R. (1998). *Beyond grammar*. Stanford, CA: Center for the Study of Language and Information.

Brown, R. & McNeill, D. (1966). The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior,* 5, 325-337.

Burzio, L. (1996). Surface constraints versus underlying representations. In J. Durand &, B. Laks (Eds.), *Current trends in phonology: Models and methods*. Salford: European Studies Research Institute, University of Salford Publications, pp. 97-112.

Bybee, J. (1994). A view of phonology from a cognitive and functional perspective. *Cognitive Linguistics*, 5, 285-305.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425-55.

Bybee, J. (1998). The emergent lexicon. *Proceedings of the Chicago Linguistic Society*, vol. 34, 421-435.

Bybee, J. L. & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language,* 58, 265-289.

Chandler, S. (1995). Non-declarative linguistics: Some neuropsychological perspectives. *Rivista di Linguistica*, 7, 233-247.

Chandler, S. (2002). Skousen's analogical approach as an exemplar-based model of categorization. In R. Skousen, D. Lonsdale, & D. B. Parkinson (Eds.), *Analogical modeling*. Amsterdam: John Benjamins, pp. 51-105.

Cole, J. (1995). The cycle in phonology. In J. A. Goldsmith (Ed.), *The handbook of phonological theory*. Cambridge, MA: Blackwell, 70-113.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Cole, J. S. & Hualde, J. I. (1998). The object of lexical acquisition: A UR-free model. In *Proceedings of the Chicago Linguistic Society*, vol. 34, 447-458.

Davis, S. (2005). Capitalistic v. militaristic: The paradigm uniformity effect reconsidered. In L. J. Downing, T. A. Hall & R. Raffelsiefen (Eds.), *Paradigms in phonological theor*. Oxford: Oxford University Press, 107-121.

Derwing, B. L. (1973). *Transformational grammar as a theory of language acquisition.* Cambridge: Cambridge University Press.

Eddington, D. (1996). The psychological status of phonological analyses. *Linguistica,* 36, 17-37.

Eddington, D. (2000). Spanish stress assignment within the analogical modeling of language. *Language*, 76, 92-109.

Eddington, D. (2002). Issues in modeling language processing analogically. *Lingua*, 114, 849-871.

Eddington, D. (2007). Flapping and other variants of /t/ in American English: Allophonic distribution without constraints, rules, or abstractions. *Cognitive Linguistics*, 18(1), 23-46.

Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S. & Dahlgren, N. L. (1993). *DARPA TIMIT: Acoustic-phonetic continuous speech corpus.* Philadelphia: Linguistic Data Consortium.

Giegerich, H. J. (1992). *English phonology: An introduction*. Cambridge: Cambridge University Press.

Goldinger, S. D. (1997). Words and Voices: Perception and production in an episodic lexicon. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing*. San Diego: Academic, pp. 33-65.

Hall, K. C. (2005). Defining phonological rules over lexical neighborhoods: Evidence from Canadian raising. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 191-199.

Harris, J. (1994). *English sound structure*. Oxford: Blackwell.

Hintzman, D. L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review*, 93, 411-428.

Hintzman, D. L. (1988). Judgements of frequency and recognition memory in a multiple-trace memory model. *Psychological Review,* 95, 528-551.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Hintzman, D. L. & G. Ludlam. (1980). Differential forgetting of prototypes and old instances: Simulation by an exemplar-based classification model. *Memory and Cognition*, 8, 378-382.

Jensen, J. T. (1993). *English phonology*. Amsterdam: John Benjamins.

Kahn, D. (1980). *Syllable-based generalizations in English phonology*. New York: Garland.

Kenstowicz, M. (1996). Base-identity constraints and uniform exponence: Alternatives to cyclicity. In J. Durand & B. Laks (Eds.), *Current trends in phonology: Models and methods*. Salford: European Studies Research Institute, University of Salford Publications, pp. 363-393.

Kiparsky, P. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, 10, 421-441.

Ladefoged, P. (2006). *A course in phonetics*, 5th edition. Boston: Thompson Wadsworth.

Lamb, S. (2000). Bidirectional processing in language and related cognitive systems. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of Language*. Stanford, CA: Center for the Study of Language and Information, 87-119.

Manelis, L. & Tharp, D. A. (1977). The processing of affixed words. *Memory and Cognition*, 5, 690-695.

McCarthy, J. (1995). Extensions of faithfulness: Rotuman revisited. Manuscript, University of Massachusetts, Amherst. ROA 110, ruccs.rutgers.edu/roa.html

McCarthy, J. & Prince, A. (1994a). The emergence of the unmarked. Manuscript, University of Massachusetts, Amherst. ROA 13, ruccs.rutgers.edu/roa.html

McCarthy, J. & Prince, A. (1994b). An overview of prosodic morphology. Manuscript, University of Massachusetts, Amherst. ROA 59, ruccs.rutgers.edu/roa.html

Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85, 207-238.

Messick, S. J., & Solley, C. M. (1957). Probability learning in children: Some exploratory studies. *The Journal of Genetic Psychology*, 90, 23-32.

Murphy, G. L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Nespor, M. & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 700-708.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 309-28.

Pawley, A. & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Smith (Eds.), *Language and Communication*. London: Longman, 191-225.

Pisoni, D. (1997). Some thoughts on 'normalization' in speech perception. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing*. San Diego: Academic, pp. 9-32.

Riehl, A. K. (2003). American English flapping: Perceptual and acoustic evidence against paradigm uniformity with phonetic features. *Working Papers of the Cornell Phonetics Laboratory*, 15, 271-337.

Rhodes, R. (1994). Flapping in American English. In *Phonologica 1992: Proceedings of the 7th International Phonology Meeting*, 217-232.

Schweitzer, A. & Möbius, B. (2004). Exemplar-based production of prosody: Evidence from segment and syllable durations. In B. Bel & I. Marlien (Eds.), *Speech Prosody 2004*. Nara, Japan: International Speech Communication Association, pp. 459-462.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations part II* . Dordrecht: Foris, pp. 337-383.

Sereno, J. A. & Jongman, A. (1997). Processing of English inflectional morphology. *Memory and Cognition*, 25, 425-37.

Skousen, R. (1989). *Analogical modeling of language*. Dordrecht: Kluwer.

Skousen, R. (1992). *Analogy and structure*. Dordrecht: Kluwer.

Skousen, R. (1995). Analogy: A non-rule alternative to neural networks. *Rivista di Linguistica*, 7, 213-232.

Skousen, R. (1998). Natural statistics in language modelling. *Journal of Quantitative Linguistics*, 5, 246-255.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Solé, M. J. (2003). Is variation encoded in phonology? In *Proceedings of the 15th International Conference on the Phonetic Sciences*, 289-292.

Stemberger, J. P. & MacWhinney, B. (1988). Are inflected forms stored in the lexicon? In M. Hammond & M. Noonan (Eds.), *Theoretical approaches to morphology*. San Diego: Academic Press, pp. 101-116.

Steriade, D. (1995). Underspecification and markedness. In J. A. Goldsmith (Ed.), *The handbook of phonological theory*. Cambridge, MA: Blackwell, pp. 114-174.

Steriade, D. (1997). Lexical conservatism and its analysis. Manuscript, University of California, Los Angeles.

Steriade, D. (1999). Lexical conservatism. In Linguistic Society of Korea (Ed.), *Linguistics in the morning calm*, vol. 4. Hanshin: Linguistic Society of Korea, pp. 157-180.

Steriade, D. (2000). Paradigm uniformity and the phonetics-phonology boundary. In M. Broe & J. Pierrehumbert (Eds.), *Papers in laboratory phonology 5*. Cambridge: Cambridge University Press, pp. 313-334.

Withgott, M. M. (1982). *Segmental evidence for phonological constituents*. Unpublished Doctoral Dissertation, University of Texas-Austin.

Zue, V. & Seneff, S. (1996). Transcription and alignment of the TIMIT database. In Hiroya Fujisaki (Ed.), *Recent research toward advanced man-machine interface through spoken language*. Amsterdam: Elsevier, pp. 464-447.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Where do Phonemes Come from? A View from the Bottom

JOHN R. TAYLOR*
University of Otago

ABSTRACT

Infants have a remarkable ability to perceive all manner of phonetic contrasts. The phonological categories of a language, however, have to be learned from experience. Two learning paradigms are contrasted – supervised learning (where learners receive feedback on their categorization attempts) and unsupervised learning (where learners rely only on properties of the input). It is argued that unsupervised learning may be the appropriate paradigm, at least for the initial stages of acquisition. Thereafter, the emergence of phoneme categories draws on various kinds of knowledge available to the learner, including knowledge of articulation, and of literacy conventions. A concluding section emphasizes the taxonomic nature of the phoneme, and suggests that the special salience of a phonemic representation reflects the status of the phoneme as a basic level category.

KEYWORDS: phoneme; perception; structuralism; categorization; unsupervised learning; basic level

* ***Address for correspondence***: John R. Taylor. University of Otago, New Zealand. Department of English (Linguistics Programme). Division of Humanities, University of Otago, PO Box 56 Dunedin, New Zealand. Phone: 64 3 479 8952, Fax 64 3 479 8558. e-mail: john.taylor@stonebow.otago.ac.nz

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I. INTRODUCTION

In a well-known passage, Saussure commented on what our mental experience would be like if we did not possess language:

> Psychologiquement, abstraction faite de son expression par les mots, notre pensée n'est qu'une masse amorphe et indistincte. … [S]ans le secours des signes, nous serions incapables de distinguer deux idées d'une façon claire et constante. Prise en elle-même, la pensée est comme une nébuleuse où rien n'est nécessairement délimité. Il n'y a pas d'idées préétablies, et rien n'est distinct avant l'apparition de la langue (Saussure, 1915: 155).[1]

Without language, Saussure claimed, thought would be inherently featureless and unstructured. For Saussure, it was language – more specifically, the conceptual categories symbolized by language – that gave structure to the amorphous substance that is prelinguistic thought. Saussure made an analogous claim about the sound substance of language. Without the mediation of a language and its phonological system, the speech signal would be equally indistinct and formless:

> La substance phonique n'est pas plus fixe ni plus rigide; ce n'est pas un moule dont la pensée doive nécessairement épouser les formes, mais une matière plastique qui se divise à son tour en parties distinctes pour fournir des signifiants dont la pensée a besoin (Saussure, 1915: 155).[2]

We can, of course, only speculate about the mental life of a person without language – a new born infant, for example, or a wild child. We are on firmer ground when it comes to our perception of the speech signal in ignorance of the linguistic categories which it encodes. When we are listening to a language which is totally unknown to us, Saussure's metaphor of the 'nébuleuse', where nothing seems clearly delineated, seems particularly apt. Learning the language consists inter alia of learning to 'make sense' of the acoustic signal, segmenting it into distinct units, classifying the units and their combinations, and, ultimately, recognizing in the signal the expression of meaningful words and phrases. This is a process which each child (with the exception, of course, of the profoundly deaf) must go through. In this paper, I comment on some aspects of this remarkable achievement, with special focus on the emergence of segmental categories.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

II. PHONOLOGICAL UNITS

One might suppose that mastering the sound system of a language would consist in learning to make progressively finer perceptual distinctions amongst the sounds that one encounters in the acoustic signal. Much ingenious experimentation, however, has demonstrated that this is not how first language acquisition proceeds. It is now well established that newborn infants are exquisitely sensitive to speech sounds, being able to discriminate all manner of contrasts which are utilized in the various languages of the world (Aslin *et al.*, 1998; Aslin *et al.*, 1981; Eimas *et al.*, 1971; Jusczyk, 1997; Kuhl, 1987; Werker & Tees, 1984). While this remarkable ability surely facilitates entry into the sound system of whatever language a child is going to learn, the ability to discriminate sounds is not sufficient for phonological acquisition to take place. A person able to perceive all manner of acoustic-phonetic differences would be rather like Luria's (1968) mnemonist, or the fictitious Funes of Borges's (1964) story – individuals with a phenomenal ability to notice and remember every detail of their experiences but who, as a consequence, are unable to generalize and form abstractions. For speech perception to get under way, it is necessary for *categories of acoustic events* to be recognized in the kaleidoscope of auditory impressions.[3] Some chunks of the acoustic signal need to be regarded, in the phonological system that is being acquired, as being 'the same' as other chunks. The first question we need to ask, therefore, concerns the nature of these chunks that the learner needs to identify. There are at least three plausible candidates with regard to the linear segmentation of the speech signal (the list is not exhaustive, and the kinds are not mutually exclusive): Words, syllables, and parts of syllables.

That competent hearers of a language perceive words in the stream of speech is self-evident. Listening to speech is essentially a matter of listening for words, and word-like units, and learning a language involves, amongst other things, learning the sound shapes of words.[4] Indeed, Jusczyk (1997: 108) suggests that the identification of words in the stream of speech is what "speech perception capacities are ultimately intended for", while others have proposed that the learner's identification of word-sized units may well bootstrap the whole language acquisition process (Beckman & Edwards, 2000; Beckman & Pierrehumbert, 2000).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

During the earliest stages of language acquisition, it may well be the case that words are learned, stored, and retrieved as phonological wholes, without internal analysis (Jusczyk, 1997; Vihman, 1996). While a reliance on gestalt storage might be viable at a time when the child's linguistic repertoire consists of at most a couple of dozen items, the increasing size of the child's lexicon necessitates other, or additional storage modalities. This is because the number of holistic sound shapes that a person could reliably differentiate and commit to memory is severely limited. As the size of the lexicon increases, some kind of internal analysis of the word-sized units becomes necessary. Thus, pieces of one word might be identified with pieces of other words, the pieces themselves might in turn be broken down into even smaller units. In this way, a relatively small inventory of phonological units, and patterns for their combination, will be able to support the learning of a large and ever expanding lexicon.[5]

Candidates for the internal analysis of words are syllables, parts of syllables (such as onsets and rhymes), and, ultimately, consonant and vowel segments. Syllables, as units of analysis, would seem to be especially appropriate for languages such as Japanese and Māori, where the number of possible syllables in the language is quite limited. This is reflected in the katakana and hiragana writings systems of Japanese, in which each syllable is represented by a distinct symbol (exactly 46 are needed.). When the number of different syllables in a language increases, internal analysis once again becomes necessary. Thus, traditional accounts of Mandarin phonology analyze the 400 or so occurring syllables (this number disregards tonal differentiation) in terms of the combination of initials and finals, i.e. onsets and rhymes. For English, and other languages with complex syllable structures, in which the number of different syllables runs into the thousands, further analysis is necessary, namely into the individual phonemes (or, perhaps better, the positional allophones) which make up the syllables.

Words, syllables, and phonemes/allophones, as units of perception and representation, all raise the same problem, namely, that of acoustic variability. A word, syllable, or phoneme can be pronounced in a virtually unlimited number of ways according to the linguistic context of the unit (its immediate phonetic environment, its place within an intonation contour, the overall rate of speech, etc.) as well as speaker-dependent properties (dialect, gender, age, speaker-specific properties of the vocal tract, and even such factors as the state of the speaker's dentures).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Bloomfield (1933) had supposed that various manifestations of a phoneme would share some common acoustic features. The invention of spectrographic analysis in the 1940's, however, and early attempts to synthesize speech by concatenating invariant segments, brought home to phoneticians in a particularly dramatic way the lack of acoustic invariance associated with the units that we hear in the speech signal (Potter *et al*., 1947). Liberman and his colleagues (Liberman *et al*., 1967; Liberman & Mattingly, 1985) developed their 'motor theory' of speech perception largely in response to this state of affairs. Specifically, they sought to locate invariance, not in the signal itself, but rather in the motor commands which gave rise to the acoustic signal. Later versions of the theory located invariance, not in the motor commands themselves, but in a speaker's "intended phonetic gestures" (Liberman & Mattingly, 1985: 2), thereby pushing the invariants into a domain which in principle is out of reach of empirical observation.

The invariance problem is a familiar one to categorization researchers. In fact, the (largely unsuccessful) search for acoustic constants in the speech signal following the invention of spectrographic analysis is merely a variation on the theme of the non-viability of 'classical' categories in general. Classical categories, it will be recalled, are defined in terms of a set of necessary and sufficient features. Especially from the 1970's onwards, it became apparent that most categories that people operate with – for example, the categories that are conventionally named by the lexemes of their language – are not in fact susceptible to classical definitions; moreover, the features which supposedly define the categories are subject to the very same problem (Taylor, 2003b, to apppear). In light of these findings, various alternative models of categorization were developed. These included prototype models (in which categories are centred around 'good examples'), probabilistic models (in which categories are defined in terms of weighted probabilities of features), and exemplar models (where categories are constituted in terms of the similarity of already encountered instances). In view of this extensive research (reviewed in Murphy, 2002; see also Mompeán, 2002: Ch. 1) it should come as no surprise that phonemes, syllables, and words should also resist definition in terms of sets of invariant acoustic features.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

III. SUPERVISED OR UNSUPERVISED LEARNING?

The 'discovery' of the phoneme has been described as "one of the most magnificent achievements of linguistic science" (Krámský, 1974: 7). The hyperbole of this statement conceals the fact that the phoneme concept is by no means a modern invention. It is the basis of all alphabetic writing systems (though, to be sure, few writing systems are consistently phonemic), and even speakers of unwritten languages are reported to have intuitive access to the phonemic structure of words.[6] Symptomatic of the popular acceptance of the notion is the fact that most monolingual and bilingual dictionaries nowadays give word pronunciations in some form of phonemic transcription. Yet, like many of the most basic concepts of linguistics – such as 'word', for example – a concise definition remains elusive, and indeed the phoneme concept has been, and remains, the subject of intense and ongoing theoretical controversies. Later in the paper I will touch on generative phonologists' rejection of the need for a distinct phonemic level of representation. In the meantime, I focus on some of the controversies which engaged the linguistic community in pre-generative days. Indeed, a glance at the journals of the time – as well as at the contents page of Joos's (1957) influential *Readings in linguistics* – gives the impression that the history of North American linguistics during the mid decades of the last century was in large part a confrontation with the problematics of the phoneme concept.

A major issue in pre-generative times concerned the criteria by which the phonemes of a language are to be established. One of the orthodoxies of the time was the prohibition on the 'mixing of levels' (Bloch, 1948; Hockett, 1942). The idea was that the investigation of a language should proceed in a strictly 'bottom-up' fashion. The investigating linguist first made detailed phonetic transcriptions of a corpus of native speaker utterances. Observation of the distribution of phonetic segments ('phones') would then permit the allocation of these segments to a fixed set of phonemes, accompanied by statements for the possible realizations of each phoneme in various contexts. Importantly, phonemic analysis was to be conducted without any reference to 'higher' levels, such as the words and morphemes of the language, nor, or course, to their meanings.[7] Subsequently, linguistic analysis would proceed to the identification of

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

allomorphs and their allocation to morphemes (again, without reference to their meaning), followed by the identification of word classes, syntactic patterns, and so on (Harris, 1951).

These discussions (for a review, see Heitner, 2005) may strike the modern reader as very arcane. Pike (1947), for one, ventured to state that no field linguist would ever proceed in the way demanded by the orthodoxy of the time, by ignoring meaning and strictly excluding any 'top-down' analysis. Nevertheless, I would suggest that the issues that were discussed in the 1940's and 1950's do relate to a matter which is very much of modern concern. Updating the discussions of more than half a century ago into more modern terminology – and fudging the distinction between the linguist's analytic procedures and the processes of language learning by children (and by machines)[8] – the question would be whether phoneme categories can emerge in unsupervised as opposed to supervised learning conditions. In supervised learning, the learner (whether human or machine) is presented with a set of stimuli which are labeled as members or non-members of the target category or categories (the labeling may take the form of feedback on the correctness or otherwise of the learner's attempts at categorization). Subsequently, the learner may be tested on new stimuli, which are presented without labeling or feedback, with the aim of determining how well the categories have been learned, and how 'ambiguous', or otherwise problematic stimuli will be handled. In unsupervised learning, on the other hand, the learner is simply presented with a set of stimuli and is required to group them into categories. The stimuli are not labeled, no feedback is provided, nor is the learner given any hints as to how many or what kinds of categories are to be formed.

It will be apparent that a strict application of the dogma of the separation of levels is in essence a prescription for unsupervised learning. Indeed, linguists of the time were much concerned with developing a set of 'discovery procedures' – that is, a set of algorithms – which would correctly, and 'automatically', identify the phonemes of a language, given only a narrow phonetic transcription. The phonemic analysis would 'emerge' from the phonetic properties of a corpus, without the analyst needing to be aware that two phonetically similar stretches were merely variant pronunciations of the same word (i.e., that the pronunciations were in free variation), or whether they in fact constituted pronunciations of different words (i.e., constituted

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

minimal pairs). If access to this latter kind of information were to be available, we would be in the domain of supervised learning.

The period in question – the 1940's and 1950's – is commonly referred to as the heyday of Bloomfieldian linguistics, reflecting the towering influence of Bloomfield's monograph *Language* (1933). It may be interesting, therefore, to recall Bloomfield's position on phonemic analysis. We have already referred to Bloomfield's belief that a sufficiently sophisticated acoustic analysis would eventually reveal the invariant properties definitional of each phoneme of a language. For Bloomfield, however, the search for these invariant properties could not be basis of phonemic analysis. Rather, linguistic analysis was based on what for Bloomfield was the "fundamental assumption of linguistics", namely, that "in every speech-community some utterances are alike in form and meaning" (1933: 78). Thus, according to Bloomfield,

> …even a perfected knowledge of acoustics will not, by itself, give us the phonemic structure a language. We shall always have to know which of the gross acoustic features are, by virtue of meanings, 'the same', and which are different for the speakers (Bloomfield, 1933: 128).

The irony of Bloomfield's position has not escaped some commentators (Harris, 1973; Taylor, 2003b). Bloomfield, who was so intent on excluding 'mentalistic' notions, such as 'meanings', from linguistic analysis, had to postulate 'sameness of meaning' as a prerequisite for any linguistic analysis at all. Be that as it may, the relevance of Bloomfield's observation to the present topic will be evident. Bloomfield was proclaiming the impossibility, in principle, of unsupervised language learning.

And, indeed, common sense would seem to be on Bloomfield's side. Consider the acquisition of word meanings. The child encounters a range of creatures of different shapes, sizes, colours, and habitats, and exhibiting different temperaments and behaviours. Some of these creatures are referred to as 'dogs', others bear different labels, such as 'cat', 'rabbit', 'cow', 'mouse', as well as 'animal' and 'pet', The child's task, now, is to work out the criteria for this classification, on the assumption that the different uses of *dog* are 'the same in meaning', that is, that they designate one and the same category of entities. It was in such terms that Brown (1958)

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

presented 'The Original Word Game'. On this account, the learner's task would be exactly analogous to that confronted by subjects participating in a supervised learning experiment. Some such process would seem to be indicated, if only because different speech communities typically categorize the environment in different ways. For example, English, French, and German distinguish 'rats' from 'mice', whereas Italian does not, both kinds of creature bearing the label 'topo'. Language-specific categories presumably do not, and could not, emerge from simple observation of the world; they have to be transmitted from one generation of speakers to the next by engagement in the 'word game'.

Appealing as it is, the 'word game', and the parallels with supervised learning, may not be the whole story. In a supervised learning experiment, a subject is presented with an array of experimental stimuli and is explicitly informed about their category membership. The counterpart of this situation in language acquisition would be that a word is explicitly associated with its referent on each occasion of its use. Yet it is not always the case that words, even words which designate easily observable entities, are uttered in the presence of their referents, and even when they are, the child still has to figure out just which features of the environment are to be matched with a given word. Gleitman (1990), in addressing the common belief that words are learned by ostension, urges us to 'look and see' whether words are indeed spoken in situations in which their referents are perceptually salient to the learners. She concludes that, in many cases, they are not. Indeed, detailed observations by Gleitman suggest that learning by ostension may actually be the exception rather than the rule. And in the case of words designating 'abstract' entities and processes, such as *think, believe*, and *know*, the words' referents may not be candidates for ostension at all. The task faced by the language learner, then, is not simply one of working out the correct categorization of an array of labeled stimuli. The learner must first discover what the stimuli are that are to be categorized.

The matter becomes more complicated still when we bear in mind that word learning is not only a question of learning semantic categories, the word forms themselves have to be learned. The learner, namely, has to realize that the multifarious ways in which *dog* can be pronounced all count as pronunciations of 'the same word'. The learner could, in principle, explore the hypothesis that variations in the duration of the vowel, or whether the final consonant

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

is released or not, might correlate with meaning differences, e.g. big dogs vs. small dogs, brown dogs vs. spotted dogs, well-behaved dogs vs. yapping dogs. Children, presumably, do not systematically explore these possible correlations between form and meaning, any more than the field linguist would test each of the myriad hypothetical senses of *gavagai*, in Quine's (1960) well-known example.[9] As Bloomfield stated, the learner would need to be apprised of the fact that the various pronunciations are, indeed, 'the same in form', as well as being 'the same in meaning'.

There are, to be sure, certain circumstances in which a learner might be explicitly alerted to the fact that different pronunciations count as 'the same', while other pronunciations are 'not the same', as, for example, when second language learners are being trained on the discrimination of minimal pairs (*ship* vs. *sheep*, and the like). The extent of this practice with children acquiring their native language is probably quite limited, and is likely to be restricted, in any case, to older children perceived to be suffering from delayed development. (We should bear in mind, also, that languages are acquired in all manner of socio-cultural settings. Whether or not children are coached in matters of pronunciation, they all – barring pathological cases – end up with adult mastery of the ambient language.) One possibility might be that learners themselves 'discover' the existence of minimal pairs, by noting, for example, that the pronunciations of *coat* refer to one kind of entity, while the pronunciations of *goat* refer to a quite different kind of entity. The need to make the conceptual distinction would therefore trigger awareness of the corresponding phonological categories. Some researchers have indeed suggested some such mechanism of phoneme acquisition (Werker & Tees, 1984).[10]

There are, however, a number of theoretical and empirical problems associated with the view that phoneme categories emerge on the back of minimal pairs. In the first place, while the existence of minimal pairs might be diagnostic of phoneme categories, it must fail as a definition of the phoneme. In English, there are scarcely any minimal pairs contrasting [ʃ], and [ʒ], or [θ] and [ð], yet we would still want to regard these sounds as belonging to different phonemes of English.[11] Moreover, the existence of minimal pairs will be largely a matter of the size of a person's lexicon. For young children, with very small vocabularies, minimal pairs, for any pair of

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

candidate sounds, are vanishingly rare. Caselli et al. (1995) list the first 50 words produced and understood by both English-speaking and Italian children. The English lists contain no minimal pairs, while the Italian lists contain only *nonna* 'granny' and *nanna* 'sleep'. Even more telling is the fact that by the age of 1, children are already well on their way to perceiving the ambient language 'phonemically' (Jusczyk, 1997), that is, they are categorizing the ambient speech sounds in line with the phonological structure of the language they are to acquire. At this stage, children have scarcely learned any words of their language at all, so cannot be relying on lexical contrasts. Once again, we are forced to the conclusion that the supervised learning paradigm – where learners have the task of categorizing labeled stimuli – simply fails to apply.

The role of supervised learning (or, rather, its absence) turns up in connection with yet another issue in language acquisition research, namely, the problem of negative evidence (Bowerman, 1988; Pinker, 1984). In supervised category learning, learners receive feedback on whether their classification of a stimulus is correct or not. Yet when it comes to the learning of the syntactic structures of their language, children are rarely given information on which of their utterances are grammatically ill-formed. Caretakers may comment on the factual correctness of a child's utterance, or on its stylistic or pragmatic appropriateness, but rarely, or not systematically, on its grammatical properties. A question that has much concerned researchers in language acquisition, therefore, is how a child comes to 'unlearn' the generalizations which give rise to utterances such as *They didn't wented*, or *He said me no*. It clearly will not do to say that the learner comes to regard these expressions as ungrammatical because they are never encountered in the input. Many things that speakers say are unique creations, never before encountered, but are not, for that reason, to be rejected as ungrammatical. One factor that seems to be involved is the child's working assumption that languages avoid synonymy (Clark, 1987). The learner comes to regard her own utterances as ill-formed to the extent that they are pre-empted by alternative wordings encountered in the input (Tomasello, 2003). Whatever the plausibility of this account, it is clear that learners must work out the properties of syntactic constructions largely on the basis of the input, its properties, and their analysis of it, not from explicit instruction or feedback on grammaticality.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The above considerations all point in the same direction, namely, that the supervised learning paradigm may not be applicable to first-language acquisition. Words do not come tagged with their semantic and phonological categories, nor is information provided on which utterances count as 'the same' in form and in meaning. I will not, in the following, pursue the question of the learning of semantic categories. With respect to phonological categories, however, there are grounds for taking seriously the reality of unsupervised learning, exactly as the structuralist insistence on the separation of levels entailed.

IV. UNSUPERVISED LEARNING OF PHONOLOGICAL CATEGORIES

Categorization has been a major research topic in cognitive psychology; for a review of the by now voluminous literature, see Murphy (2002). Surveying this literature, one is struck by the fact that the vast bulk of the research has been in the supervised learning tradition, employing procedures that in the psychological literature are commonly referred to as 'category formation' experiments. The term may actually be something of a misnomer, since the categories in question have already been formed, namely, by the experimenter; the subject's task would therefore be more accurately described as one of problem solving rather than category formation (Fodor 1980). The subject, that is, has to work out the criteria by which certain stimuli have been put into a certain category, whereas other stimuli have not. Much of this research has been conducted on the example of visually presented stimuli; in comparison, the categorization of (non-linguistic) auditory stimuli has been neglected (but see Lotto, 2002). There is, however, a modest tradition of concept formation experiments conducted on the example of phonological categories (Jaeger, 1980, 1986; Jaeger & Ohala, 1984; Mompeán, 2002; Weitzman, 1992).

As mentioned, surprisingly little research has been conducted by cognitive psychologists on unsupervised learning, or 'category construction', as Murphy (2002: 126) calls it, in contradistinction to 'category formation'. The little research that Murphy reports suggests that the categories that subjects spontaneously construct in such experiments are quite different from the categories that they normally operate with. There is a tendency, namely, for subjects to seize on a single dimension of the stimuli, such as their size, or colour, and to group them accordingly

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

(Murphy, 2002: 128). The complex, multi-dimensional, and probabilistic categories enshrined in the lexicons of human languages rarely emerge.

Further perspectives on supervised and unsupervised learning are provided by the computational modelling of learning, especially in artificial neural networks (McLeod *et al*., 1998). (It is, in fact, from the computational literature that I have taken the terms 'supervised' and 'unsupervised'). Consider a typical connectionist set-up. An array of input nodes is linked, possibly via one or more sets of hidden nodes, with an output array. Initially, the nodes are connected by randomly assigned connection weights. An input is presented, and the system's output is compared with the 'desired' output. The connection weights are then adjusted so as to decrease the system's error. The cycle is repeated – typically, many thousands of times – with each input being matched with a desired output. Eventually, the connection weights stabilize, and the system may be able to give the 'correct' output even for new inputs which it has never before encountered. One of the earliest and best-known applications of this procedure to language learning is Rumelhart & McClelland's (1986) account of the training of a network to produce past tense forms of English verbs (for an update, see Plunkett, 1995). The procedure, it will be appreciated, rather closely models the psychologists' category formation experiments. Thus in the psychologists' experiments, we might suppose that at first subjects cannot make head or tail of the array of stimuli that they are presented with, and, like the neural network, give random responses. After repeated trials, in which feedback is provided, they increasingly come up with the 'correct' classification of the stimuli.

Unsupervised learning in artificial neural networks involves the automatic recognition of patterns and regularities in the input. Several aspects of linguistic structure have been subjected to this kind of procedure. Thus, Goldsmith (2001) proposes a heuristic for automatic morpheme segmentation, while other aspects of linguistic analysis are addressed in Broeder & Murre (2000). For example, for Gillis, Daelmans, and Durieux (2000), the issue is the learnability of word stress rules on the basis of syllable structure and segmental features, while for Shillcock et al. (2000) the problem is to identify words from a phonemic transcription of connected speech (from which, of course, the word spaces had been removed). A common technique in unsupervised learning involves the use of clustering algorithms (Manning & Schütze, 1999). Each stimulus is defined as

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

a point in multi-dimensional space, and inputs cluster according to their relative closeness, with each new stimuli being 'categorized' in terms of the cluster it gets associated with. One of the best-known unsupervised procedures is the self-organizing maps of Kohonen (1982); a more sophisticated model has been developed by Kasabov (2002). Employing Kasabov's ECOS (= 'Evolving Connectionist Systems') model, Morales & Taylor (2005) found that the unsupervised learning of small vocabularies, the sole input to the system being digitized 'signatures' of different pronunciations of the words, turned out to be remarkably robust at the testing phrase, that is, in correctly classifying new pronunciations of the words.

Under what circumstances might unsupervised learning take place in human subjects? One condition would be that the stimuli naturally cluster into so many categories. It might be the case, for example, that different sets of features co-occur in distinct sets of stimuli, or that a continuously varying feature has frequency-of-occurrence values that are bi-modally distributed over the stimuli. In such cases, the categories might be said to be 'in the world', in that the relevant categories can be identified in terms of feature correlations or feature maxima.

It goes without saying that the learner has to be able to perceive the features in question. Consider, in addition, the possibility the learner may be innately predisposed to respond to certain features, or to certain dimensions of the stimuli. In this case, the emerging categories would be a function of the system's perceptual mechanism, rather than feature correlation in the world. We can illustrate the issues on the categorization of colour. On the one hand, it could be argued that the colour solid represents a three-dimensional array of all possible colours (the three dimensions being hue, saturation, and brightness), with no natural boundaries or lines of segmentation. The colour solid does not naturally divide into so many categories. This aspect must be counterbalanced by the fact that not all the possible colours occur equally frequently in the environment. Regions in the colour space which dominate in the environment might therefore be good candidates for emergent categories. Research into the linguistic encoding of colour, however, has shown that different languages around the world tend to select their colour categories from a universal set of focal colours (Berlin & Kay, 1969). The focal colours are those which the human visual system is specifically attuned to respond to, such as red and green, blue and yellow in the first instance, and admixtures of these, such as orange, pink, and so on. The

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

focal colours are the ones that tend to be lexicalized first, in spite of the fact that they may occur relatively infrequently in nature (Taylor, 2003b)

In light of the above remarks, let us now return to the learning of phonological categories. One of the most intensively studied features of the acoustic-phonetic signal is the role of Voice Onset Time (VOT) in the differentiation of different kinds of stops, such as voiced vs. voiceless, unaspirated vs. aspirated (Liberman *et al*., 1958; Lisker, 1978).[12] It has also been established that prelinguistic infants are highly sensitive to differences in the VOT continuum (Eimas *et al*., 1971). Some scholars, including Eimas, have suggested that this fact alone may be sufficient to trigger the formation of the respective categories; there would, therefore, be grounds to claim that the categorization of stop consonants is driven by innate properties of the human perceptual mechanism. Complicating the situation, however, is the fact that different languages exploit the VOT dimension in different ways. To the extent that VOT defines language-specific categories, these categories presumably have to be learned from experience. But even within a single language, it may be inappropriate to refer to *the* VOT values which differentiate the different categories of stops. VOT depends on many variables, such as the place of articulation of the stop (VOT values for bilabials are, on the whole, shorter than for velars; Lisker & Abramson, 1964), the prosodic properties of the syllable, i.e. whether stressed and foot-initial, or unstressed, the overall speech rate, and whether in utterance-initial position, and so on. These variations are subtle and numerous, and native proficiency in a language requires that they be learned (Pierrehumbert, 2003).

Leaving aside these various sources of variation, let us consider the simplified case of stops in syllable- and foot- initial position, that is, in the onset position of stressed syllables, the kind of sounds, namely, that have been so intensively studied in the experimental literature over the past decades. Imagine two hypothetical languages, in which foot-initial VOT values between, say, -50 and +50ms, occur with more or less equal frequency. One language places the boundary between voiced and voiceless stops around +5 ms, the other places the boundary between voiceless unaspirated and voiceless aspirated stops around +25 ms. It will be apparent that the unsupervised learning of the respective categories will be all but impossible. The learner would need the information that in the one language, VOT-values of +10 and +40 count as 'the same',

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

in the other language they count as 'different'. Unsupervised learning, however, *would* be feasible, if VOT values were bimodally distributed, clustering, for example, around +5 and +30. Frequency distribution of the stimuli would therefore naturally divide the stimuli into two categories. As it happens, VOT values in natural languages (in a given prosodic position) do indeed tend to be distributed in this way (Lisker & Abramson, 1964).

The possibility that distributional properties of the input might drive category formation was investigated by Kornai (1998) with respect to the vowel formant data reported in Peterson & Barney (1952). Peterson & Barney took measurements of the first three formants of 10 American English vowels each of which was spoken twice by 76 talkers (men, women, and children). A first glance at a graph plotting the formant data for all the vowel tokens gives the impression of a broad swath of values, with few natural boundaries. Even so, as Kornai observes, the formant data present a picture very different from a random set of dots in 2- or 3-dimensional space. Kornai reports, in fact, that automatic clustering procedures were able to assign the formant values to 10 categories, whose central values corresponded rather closely with those of the 10 intended vowels.[13]

Could first language learners exploit distributional facts in the input to bootstrap the learning of phoneme categories? There is some evidence to suggest that they could. With respect to their sensitivity to statistical properties of the input, it has been demonstrated that prelinguistic children, when presented with strings of nonsense syllables, are able to utilize statistical information in order to identify recurring patterns of syllables as 'words' (Saffran *et al*., 1996). Support also comes from Maye & Gerken (2000) and Maye et al. (2002), who exposed learners to a range of stop-vowel stimuli which, in one condition, were unimodally distributed in terms of their frequency of presentation, and, in the other condition, bimodally distributed.[14] When tested, learners in the latter condition (infants as well as adults) responded in a way suggesting that they had constructed two categories, whereas learners on the first condition did not. As Maye & Gerken (2000: 530) remark, it is as if listeners "maintain some sort of mental histogram", tracking the frequency of occurrence of acoustic patterns they had encountered. Anderson et al. (2003) make a similar point, hypothesizing that the sequence in which phonological categories are acquired is driven by input frequencies.[15]

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

It is therefore entirely plausible that the phonetic categories distinctive of a particular language (such as the aspirated vs. unaspirated stops, or the various vowel categories) could be 'seeded' during the first years of life by the statistical properties of the input. Further exposure to the language will, of course, be needed in order to sharpen and refine these categories (Bohn 2000). There is evidence that this process may continue until well into the school years (Hazan & Barrett, 2000). Some additional aspects of this process are mentioned below.

(i) Although a single dimension (such as VOT for the stop consonants, or formant frequencies for the vowels) may be sufficient to seed the respective categories, further exposure may enrich the category representations through the accretion of correlated properties. While VOT has been shown to be reliable cue for different kinds of stop consonants, VOT is not the only dimension differentiating the syllable-onset stops in English (Lisker, 1978). The intensity and spectral properties of the burst, the rate of change of formant transitions, and even the pitch of the ensuing vowel tend to correlate with the aspirated/unaspirated distinction, thus providing additional, though possibly redundant, cues for the characterization and differentiation of the respective categories. For vowels, an additional differentiating aspect is variations in duration (Peterson & Lehiste, 1960), and even differences in inherent pitch. Thus, all other things being equal, the duration of the vowel in *sad* [sæd] is likely to be greater than the duration of the vowel in *said* [sɛd].[16]

(ii) As acquisition progresses, the categories will become subject to internal organization. Members of the same category will come to be perceived as increasingly similar, while perceptual differences between neighbouring categories are increased. Kuhl (1991) in this context speaks of the 'perceptual magnet effect' – outlying members of a category tend to be 'drawn in' towards its prototypical centre. Thus, speakers become increasingly desensitized to differences between stimuli belonging to the same category, but readily discriminate stimuli which lie just on either side of a category boundary. These constitute the well-studied phenomenon of categorical perception, defined, by Harnad (2003) as a situation where "perceived within-category

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

differences are compressed and/or between-category differences are separated, relative to some baseline of comparison".

(iii) The increasing size of the learner's lexicon may also be a factor in phonological development (Beckman & Edwards, 2000). One might suppose that the ability to discriminate categories of sounds will entail that learners integrate these categories into their mental representations of words. Pater et al. (2004), however, report that infants who are able to discriminate *pin* and *bin*, *bin* and *din*, etc., were unable to associate these syllables with meaning differences in a word learning experiment. They explain this seemingly paradoxical finding in terms of the additional processing demands of word learning, involving the association of the acoustic stimuli with referential meaning. In early stages of language acquisition, therefore, words may well be represented in terms of their gross acoustic properties. As word learning gets under way, and the child's lexicon increases in size, more accurate lexical storage will be necessary. This will not only strengthen the mental representation of the phonological categories, it will also reinforce their differentiating potential.

(iv) A further factor in the acquisition of phonological categories is the various 'knowledge effects' that come into play. I address this issue in the next section.

V. KNOWLEDGE EFFECTS

I have given a tentative account of how the phonetic shapes of words, such as *coat* and *goat*, *ship* and *sheep*, might plausibly be learned in an unsupervised learning situation. This account, however, does not equate to the learning of phonemes, as these are traditionally understood. What our hypothetical learner will have acquired are allophones, or "phonetic equivalence categories" (Maye & Gerken, 2000: 532), that is, categories which comprise sounds which occur in particular phonological positions. The key characteristic of phonemes is "equivalence across contexts" (Pierrehumbert, 2003: 118). The phoneme, namely, is the level of representation at which *coat* and *goat*, *lack* and *lag*, *anchor* and *anger*, *bicker* and *bigger* differ with respect to the very same

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

contrast, namely, /k/ vs. /g/. The words contain the 'same' sounds, albeit in different syllable and prosodic positions. Unsupervised learning might result in the acquisition of the properties of each of the above words, yet not deliver the insight that *coat* and *goat* differ in the same way as *lack* and *lag*. On what basis, therefore, can we say that *coat*, *lack*, *anchor*, and *bicker* all contain the 'same sound', namely /k/?

The standard structuralist answer to this question was that different sounds belong to a single phoneme category because of their similarity and their interchangeability. Referring to Z. Harris's (1951: 20) statement that "[i]t is empirically discoverable that in all languages which have been described we can find some part of one utterance which will be similar to a part of some other utterance", Hoijer (1958, cited in Heitner, 2005: 20) comments:

> 'Similar' here means not physically identical but substitutable without obtaining a change in response from the native speakers who hear the utterance before and after the substitution: e.g., the last part of 'He's in' is substitutable for the last part of 'That's my pin' (Hoijer, 1958: 573).

Drawing on the structuralist tradition, Quine (1987: 150) gave the following account:

> Two distinguishable sounds belong to the same phoneme, for a given language, if switching them does not change the meaning of any expression in that language: such is the ordinary uncritical definition of the phoneme (Quine, 1987: 150).

Quine immediately modifies this in an attempt to exclude the controversial reference to 'meaning':

> But meaning is a frail reed; surely the phonemes, the very building blocks of the language, are firmer than that. They are indeed, despite occasional misgivings to the point. There is an easy behavioral criterion of sameness of phoneme that presupposes no general notion of sameness of meaning. Two sounds belong to the same phoneme if substitution of one for the other does not affect a speaker's disposition to assent to any sentence (pp. 150-151).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The claims made here are open to question on several counts. Consider, first, the issue of substitutability. Just as Pike (1947) queried whether linguists of his time really did pursue phonemic analysis without any reference to meaning, we can also ask ourselves whether anybody ever did perform the substitution tests. Nowadays, since the advent of digital signal processing, it is a relatively simple matter to cross-splice parts of recorded utterances. In earlier times, the experiment would have involved (literally) cutting up lengths of magnetic tape and sticking the bits together in a different sequence – a messy and time-consuming process at best, and prone to all kinds of errors and misjudgements. If such substitution experiments had been performed, the responses of native speakers might not at all have corroborated the phonemic analyses that the investigator was trying to validate. For example, if one cross-splices the initial /h/ sounds of *who* and *heat*, the resulting forms do not at all sound like *who* and *heat*, or even like English words at all. Or consider the initial and final consonants of a word like *tot*. If the final 't' were to be glottalized – a rather frequent pronunciation in many accents – interchanging the initial and final segments would, if anything, produce a word roughly transcribable as [ʔɒt^{h}], and heard as something like *ott*. Again, the two 't's cannot reasonably be said to be substitutable. When linguists, whether professional like Hoijer, or amateur like Quine, made statements about 'substitution' as the basis of phonemic analysis, we are dealing, I suspect, with armchair experimentation, intended to give a spurious air of scientific grounding to the enterprise.[17]

The claim that sounds are assigned to the same phoneme category on the basis of their phonetic similarity also does not hold up to scrutiny. As mentioned, the initial segments of *who* and *heat* – which, in terms of their articulation, are voiceless anticipations of the following vowels, phonetically [u̥] and [i̥] – do not sound at all similar, when excised from their context. (One would not, for example, want to claim that whispered versions of [u] and [i] are 'similar', and for this reason assign them to the same phoneme category). Likewise, there is little acoustic similarity between a glottal stop and an aspirated [t^{h}], in the above-mentioned pronunciation of *tot*. Conversely, the unstressed [ɪ] in *classify*, if voiceless (which it might well be), is essentially the same sound as the initial /h/ of *hit* (Pierrehumbert, 2003: 129), yet no one, presumably, would want to claim that the sounds are members of the same phoneme.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

We need to look elsewhere for the source of intuitions about "equivalence across contexts". The place to look, I suggest, is the 'knowledge base' of categories (cf. Mompeán, 2004). In a seminal paper, Murphy & Medin (1985) posed the question why some groupings of objects are "informative, useful, and efficient", whereas others are "vague, absurd, or useless" (p. 289). They dismiss as simplistic the view that entities cohere in a category on the basis of their similarity; after all, some kind of similarity can be perceived in any grouping of objects. Rather, category coherence is a function of some 'underlying principle', or 'theory', which 'explains' why the entities should be grouped together, e.g. in terms of encyclopaedic knowledge of the domain, presumptions about causal connections, or the role of the entities within scripts and scenarios. Rather than entities being categorized on the basis of their similarity, it is the 'theory' relating the entities that makes them seem similar (p. 291). The intuitions of native speakers (and of linguists) that *who* and *heat* begin with similar-sounding segments would be the consequence of phonemic categorization, not its cause.

With regard to phonetic segments, an important piece of knowledge concerns how these sounds are made – which articulators are involved, manner of airflow, and so forth. Thus, for English, the initial and final segments of *tot* both involve alveolar closure with no accompanying vocal fold vibration, even though the acoustic effects of the articulation are very different for onset and coda consonants. Knowledge of the articulation could therefore support the grouping of the initial and final consonants into a single category. Jusczyk, in fact, has argued that a major impetus for the emergence of phoneme categories could well be the need for the learner to coordinate perception and production:[18]

> From the standpoint of word recognition, there is no need of an ability to detect the similarity in the initial portions of the words "big," "beet," "bop," and "bun." Nor is there any particular need for the speech perception system to extract any similarity between the way that the word "park" begins and the way that "tip" ends (although this ability is critical for learning to read English). However, in order to produce, and reproduce, any of these items correctly on another occasion, it may be helpful to take note of any similarities in the articulatory gestures that are required to produce these (Jusczyk, 1997: 205).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

These remarks are relevant also for non-canonical articulations of stops. For example, stops in other than onset position of stressed syllables might not achieve full closure. *Upper* may be pronounced as [ˈʌɸə], *bigger* as [ˈbɪɤə]. Here, the gesture towards closure is made, but is not fully achieved. Knowledge of the articulations may thus support the intuition that [ɸ] and [ɤ] are members of the /p/ and /g/ categories, respectively. Or consider the fact that in many accents of English, a coda /t/ (under certain prosodic conditions) is typically glottalized, that is, the alveolar closure is made simultaneously with a glottal closure [ʔ͡t] If, furthermore, glottal closure should momentarily precede alveolar closure, there may well be no trace of the alveolar closure in the acoustic signal. The alveolar gesture may nevertheless be present, and could give rise to the intuition that the final [ʔ] is still a kind of /t/. Examples of this kind of 'hidden' articulation (that is, articulations with no auditory consequences) are documented in Browman & Goldstein (1992).

A second source of knowledge concerns dialectal and stylistic variants. There is no such thing as the perfectly homogeneous linguistic community of Chomsky's (1965: 1) idealization. Even leaving aside dialectal variation, each speaker commands a range of stylistic varieties, and comes into contact with many different speaking styles. Observing that *cat* is variably pronounced [kæt^{s}], [kæt^{h}], [kæt$^{\neg}$], [kæʔ͡t], and [kæʔ], the learner may come to group these different coda sounds as different kinds of /t/. Similarly, observing that in slow, careful pronunciation, *city* has a medial [t], whereas in rapid speech it has the flap [ɾ], the flap may again be assimilated to the /t/ category, in spite of its phonetic distinctiveness.

A third influence would be knowledge of the orthography. The flap in *city* might well be identical in articulation to the flap in *ready*. Knowledge of how these words are spelled, however, could cause the first to be categorized as a kind of 't', the latter as a kind of 'd'. Knowledge of morphological relations might also come into play. The perception of the flap in *madder* as being an example of /d/ rather than /t/ could be a consequence of the fact that the speaker knows that *madder* is derived from *mad*. Both these issues are extensively discussed in Mompeán (2004).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I have outlined some of the factors which might contribute to the emergence of the phoneme concept. The possibility still remains, however, that phoneme categories might be inventions of analyzing linguists, which play no role in the mental representations of linguistically naïve speakers. As Jaeger (1980: 233) put it, "even the most basic or self-evident" claims of theoretical linguistics need to be subjected to empirical investigation. A number of scholars, including Jaeger (1980) and Mompeán (2002: Experiment 3), have indeed attempted to demonstrate the psychological reality of the phoneme concept, with encouraging results. Thus, Jaeger found evidence that subjects classified various allophones and positional variants of /k/ into a single category, while Mompeán reported analogous findings for the allophones of /p/. It should be borne in mind, however, that both researchers employed a concept formation paradigm,[19] in the supervised learning tradition, as described above. It cannot, therefore, be ruled out that the experimental subjects were simply able to solve the categorization puzzle that they had been presented with, with no implications that the subjects had prior mental representations of categories, nor, even less, that the categories played a role in the subjects' day-to-day linguistic performance.[20] On the other hand, the fact that all 9 of Jaeger's subjects, and all 20 of Mompeán's, were able to form the categories to criterion,[21] would suggest that the subjects were indeed tapping into their mental representations of the respective categories, rather than constructing ad hoc categories in response to the experimental tasks.

A word of caution is necessary, however. English speakers who 'have' the relevant phoneme categories readily appreciate that *cat*, *tack*, and *act* contain the 'same' three sounds, arranged in different sequences. Indeed, the insight that words can be segmented into smaller units, and that these units recur in different words, would seem to be a prerequisite for mastery of an alphabetic writing system (Treiman & Baron, 1981), even though, as the example of *cat, tack*, and *act* shows, the correspondence between phonemes and letters is not always one-to-one. Continuing experience with an alphabetic writing system will only serve to strengthen and entrench the phoneme concept and its application to the words on the language. As Kornai (1998) observes, "to the extent that a … phoneme based writing system can easily be acquired and consistently used by any speaker of the language, the psychological reality of the units forming

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

the basis of the system becomes hard to deny". To be sure, the segmental phonemic analyses made by children learning the writing system may not always correspond with the analyses of professional linguists, nor even with the analyses that are enshrined in the writing system (Treiman, 1985). Neither is it the case that all children achieve the phonemic insight at the same time, and at the same rate. Even 'wrong' segmentation, though, still testifies to the implementation of some segmentation strategy, and could be taken as evidence that the phonemic insight is present. It may be, however, that some speakers *never* achieve the initial phonemic insight. Mattingly (1972) and Sampson (1985: 163) suggest that residual levels of illiteracy, even in societies with universal education, are due to the fact that a small minority of individuals fail to appreciate the phonemic structure of their language. While illiteracy, in a predominantly literate society, obviously impacts on a person's linguistic development in many ways (for example, by depriving them of exposure to literary styles and genres, and their associated syntactic and lexical properties), we probably should not conclude that the basic speaking and listening abilities of these individuals will be substantially impaired vis-à-vis that of their literate compatriots. Maye & Gerken (2000: 532) suggest that "phonetic equivalence categories … could plausibly be the *only* psychological correlates" (authors' emphasis) of the linguist's phonemes. In view of the experimental evidence cited above, as well as the fact of widespread literacy in alphabetic writing systems, this is probably an overly cautious view. On the other hand, knowledge of phonetic equivalence categories (i.e. positional allophones) could plausibly be *sufficient* for speaking and listening proficiency to be guaranteed.

VI. CONCLUDING REMARKS

As stated earlier in this paper, the phoneme concept is controversial. I have framed the above discussion around some of the controversies which were current during the heyday of Bloomfieldian structuralism, in the mid decades of the last century, concerning the criteria by which phonemes are to be identified. As is well-known, the advent of generative phonology, in the 1960's and 1970's, ushered in new controversies. Specifically, generative phonologists such as Postal (1968) and Chomsky & Halle (1968) queried the need for a phonemic level of

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

representation at all, proposing instead that a battery of ordered rules was able to generate the surface form of an utterance (roughly, the utterance in a narrow phonetic transcription) directly from a unique representation of each constituent morpheme, with no special theoretical status attaching to an intervening phonemic representation. Crucial to their argument for ignoring the phoneme was the fact that certain rules (e.g. of assimilation) sometime seemed to bypass the phonemic level altogether,[22] while others appeared to create surface contrasts (i.e., minimal pairs) which did not correspond with intuitions about a phonemic level.[23] Even so, as Schane (1971) observed, the output of one set of rules – the morphophonemic rules – did correspond, by and large, with what would earlier have been called a phonemic representation, while the phonetic rules corresponded, by and large, with what would have been regarded as phoneme realization rules.

The generative phonology approach, it will be observed, was strictly 'top-down', in the sense that details of surface pronunciations were derived from more abstract representations, rather than vice-versa. Generative phonology thus inverted the 'bottom-up' programme of the Bloomfieldians. From the perspective of the child acquiring an ambient language, a top-down approach can be viable only if one makes the gratuitous assumption that the abstract units are already available to the learner, namely, through genetic inheritance (Lindblom, 2000). This is a dubious proposition, if only because of the language-specificity of the more abstract categories (such as the phonemes). If we make – as I think we should – minimal assumptions concerning the learner's initial state, we are obliged to consider seriously the bottom-up perspective. This has been my aim in this paper.

Chomsky (1964) spoke disparagingly of the 'taxonomic phoneme'. Underlying the present account is, on the contrary, the view that phonemes are properly regarded as categories whose members are positional allophones; these latter in turn are also categories, whose members are encountered utterance events (cf. Nathan, 1986). Phonemes, therefore, take their place within a taxonomy of phonetic segments. The taxonomy need not, of course, stop at the phoneme. Phonemes might be grouped together in higher level, i.e. more 'schematic' categories, such as 'vowel' and 'consonant', with several intervening categories in between, such as 'obstruent', 'nasal', 'front vowel', 'short vowel', and so on. The very essence of the phoneme is, therefore, its

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

taxonomic status. If this view of the phoneme is accepted, then the extensive research on categorization and taxonomies that has been conducted by psychologists, and cognitive scientists more generally, becomes relevant to phonological theory. It becomes legitimate to enquire, for example, whether phoneme categories exhibit prototype effects, whether phonemes might be considered as 'basic level' categories within a taxonomic hierarchy, and what the distinctive properties of categories that are 'superordinate' and 'subordinate' to the basic level might be. These questions are touched on in Taylor (2002: to appear). A full discussion, however, must await a sequel to this paper.

NOTES

1. "Psychologically, setting aside its expression in words, our thought is simply a vague, shapeless mass. ... [W]ere it not for signs, we should be incapable of differentiating any two ideas in a clear and constant way. In itself, thought is like a swirling cloud, where no shape is intrinsically determinate. No ideas are established in advance, and nothing is distinct, before the introduction of linguistic structure." (Saussure/Harris, 1983: 155)

2. "The substance of sound is no more fixed or rigid than that of thought. It does not offer a ready-made mould, with shapes that thought must inevitably conform to. It is a malleable material which can be fashioned into separate parts in order to supply the signals which thought is in need of." (Saussure/Harris,1983: 155).

3. Strictly speaking, of course, the input to acquisition is not just auditory, but (in the case of sighted individuals) auditory-visual, in that the learner has access to visual information pertaining to the speaker's lip and jaw movements.

4. This statement leaves open the question of what constitutes a 'word' for purposes of perception, storage, and retrieval. The category comprises, in the first instance, word forms, such as *run, runs, running*, but also 'phonological words', such as *cuppa* [kʌpə] in *cup of tea*. Frequently occurring combinations, such as *all gone, bye-bye*, and *good-night*, might also have word-like status, at least for the young child. These issues, though important, are not strictly relevant to the point made in this paragraph.

5. Analysis and segmentation does not, of course, entail that words will cease to be stored as wholes. To claim this, would be to fall foul of the 'rule-list fallacy' (Langacker, 1987). It is plausible, indeed likely, that words continue to be stored as phonological wholes at the same time as their phonological constituents are recognized (cf. Lachs *et al*., 2000).

6. Cf. Sapir's (1921: 56) often-cited remark: "In watching my Nootka interpreter write his language, I often had the curious feeling that he was transcribing an ideal flow of phonetic elements which he heard, inadequately from a purely objective standpoint, as the intention of the actual rumble of speech".

7. A couple of representative statements: Hockett (1942: 20) asserted that "no grammatical fact of any kind is used in making phonological analysis", while Bloch (1948: 5) declared: "we shall avoid all semantic and psychological criteria. The implication is, of course, that such criteria play no part, or at least need not play one, in the theoretical foundation of phonemics. … The basic assumptions that underlie phonemics, we believe, can be stated without any

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

mention of mind and meaning". Bloch did, however, concede that in practice a linguist *would* appeal to meaning, but only as a "shortcut".

8. For the Bloomfieldian structuralists, with their fiercely anti-mentalist stance, it would have been unscientific to attribute psychological reality to their analyses. Nowadays, largely as a result of the 'cognitive turn' initiated by Chomsky, we have few qualms about crossing the boundary between the subject matter of linguistics and the cognitive states and processes of a language user. A linguistic description, namely, is taken as a hypothesis about a speaker's mental representations, and a linguist's analytical procedures may be regarded as analogous to those of the child language learner.

9. Quine (1960) posed the question of how a field linguist, on observing the native to utter *gavagai* on seeing a rabbit, would establish that *gavagai* means 'rabbit'. *Gavagai* could mean many things, only some of which could be subject to empirical disconfirmation.

10. See, however, Werker (2003) for a more recent view of the matter.

11. Daniel Jones insisted that "it is incumbent on us to distinguish between what phonemes *are* and what they *do*" (Jones, 1973 [1957]: 28). Thus, the possibility of lexical contrasts (i.e. minimal pairs) should be seen as a consequence of the existence of phonemic categories, not their defining, or causal feature: "An important point to notice is that the phoneme is essentially a phonetic conception. The fact that certain sounds are used in a language for distinguishing the meanings of words doesn't enter into the definition of a phoneme. It would indeed be possible to group the sounds of a language into phonemes without knowing the meaning of any words" (Jones, 1929, quoted in Bloch, 1948: 6). For a critique of the view that phonemes are inherently contrastive entities, see Berg (1993).

12. VOT is the duration, usually measured in milliseconds, between the release of a stop closure and the onset of voicing, typically diagnosed by the presence of periodicity in the wave form. A positive VOT value, e.g. +50, indicates that voicing sets in after the release; a negative value, e.g. -50, indicates that voicing sets in before the release.

13. The clustering experiment reported by Kornai did, however, specify the number of target clusters as 10. This would correspond to a situation in which a learner is informed about the number of vowel categories in a language, and is left to work out to which of the categories individual tokens are to be assigned.

14. The experiments were conducted with English speakers, and concerned the contrast between initial voiced stops (as in *die*) and unaspirated voiceless stops, such as occur, after an initial 's', in *sty*. The contrast is alien to the phonological system of English.

15. A reviewer takes issue with the notion of the statistical learning of phonetic categories, pointing out that "speakers are not tape recorders... they don't just record sound images and compare them." However, the results obtained by Maye and Gerken (2002) are very strong evidence that listeners do indeed record and compare even minute phonetic details of heard utterances; without some such mechanism, it is difficult to imagine how their results could be explained at all. Listeners' attention to, and retention of, fine acoustic-phonetic detail is also supported by research by Goldinger (1996) and by Lachs, McMichael and Pisoni (2000). Circumstantial evidence is the fact, noted by Pierrehumbert (2003, 120), that the properties of phonetic categories are in the main language-specific; consequently, these properties "must be learned by native speakers, because they have consequences for category boundaries in perception and because they must be accurately reproduced to achieve a native accent in production." It may be relevant, also, to recall that for Slobin (1985), one of the "operating principles" enabling language acquisition to take place, was: "Keep track of the frequency of occurrence of every unit and pattern that you store". The role of the frequency of occurrence in language acquisition has been reviewed by Ellis (2002).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

16. Bohn and Flege (1990; 1992) report that second language learners of English – for whom the [æ]–[ɛ] and [i]–[ɪ] contrasts are notoriously difficult – may rely predominantly on durational differences, unlike native English speakers, who rely predominantly on spectral differences.

17. A more generous interpretation of the substitution test would be that the analyzing linguist would replace one phonetic symbol in a transcription by another and then try to articulate the result. The feasibility of this enterprise presupposes that phonetic symbols accurately represent the acoustic properties of the speech signal, which, in the case of the stop consonants, is at best questionable.

18. Barring pathological cases, speakers are also listeners, and vice versa. Thus, a speaker is continually presented with the auditory consequences of her own articulations. Without needing to subscribe to the now largely discredited motor theory, with its claim that speech sounds are perceived in terms of the articulations that produced them, we can suppose (as a reviewer has suggested) that listeners will be inclined to intuit the articulatory intentions of a speaker. This suggestion links up with a wider theme in the acquisition literature, namely, the view that language acquisition may be driven by the learner's ability to read the intentions of an interlocutor (Taylor, 2002: 67-8; Tomasello, 1999). The matter has been investigated mainly from the point of view of the learning of word meanings, rather than with respect to the learning of phonetic categories.

19 This, at least, is true of Jaeger's Experiment 2. Experiment 1 used a classical conditioning paradigm, in which the results from 10 out of 16 subjects had to be discarded.

20. Imagine, for example, a concept formation experiment, in which the concept to be acquired is defined by the features [two-syllable word] and [beginning with either /l/ or /s/]. The fact that some subjects might be able to form this category to criterion would not entitle us to infer that the category plays any role whatsoever in the subjects' mental representation of their language.

21. In contrast, 6 of Mompeán's 20 subjects (2002: Experiment 1) failed to form the category '(word-initial) consonant'. This finding could be interpreted to mean that the superordinate category 'consonant' is less available to consciousness than a basic level phoneme category such as /p/.

22. For example, nasal consonants typically assimilate to the place of articulation of a following obstruent. In words such as *link* [lɪŋk], *imp* [ɪmp], and *sent* [sɛnt], the assimilated nasals would be assigned, unproblematically, to the phonemes /ŋ/, /m/, and /n/, respectively. But in the case of *comfort* [kʌmfət] and *camphor* [kæmfə], it is by no means obvious to which phoneme the assimilated [m] should be assigned. In the first set of examples, assimilation determines the occurrence of different phonemes, in the second set, assimilation results in a sound whose phonemic status is uncertain. The two sets can be unified by assuming an underlying nasal segment, which receives its place feature through assimilation, thereby removing the need for a distinctive phonemic level of representation.

23 An example of this kind of spurious minimal pair is the contrast, in some dialects, between *cat* [kæt] and *can't* [kæ̃t].

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

REFERENCES

Anderson, J.. Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech*, 46, 155-182.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development*, 52, 1135-1145.

Aslin, R. N., Jusczyk, P. W., & Pisoni, D. B. (1988). Speech and auditory processing during infancy: Constraints on and precursors to language. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and aanguage*, vol. 2. New York: Wiley, pp. 147-254.

Beckman, M. & J. Edwards. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development*, 71, 240-249.

Beckman, M. & J. Pierrehumbert. (2000). Positions, probabilities, and levels of categorisation. Keynote address, *Eighth Australian International Conference on Speech Science and Technology*, Canberra, Dec. 4-7, 2000. Available at http://babel.ling.northwestern.edu/~jbp/SST2000.pdf

Berg, Th. (1993). The phoneme through a psycholinguist's looking-glass. *Theoretical Linguistic*s, 19, 39-76.

Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.

Bloch, B. (1946). A set of postulates for phonemic analysis. *Language*, 24, 3-46.

Bloomfield, J. (1933). *Language.* London: George Allen & Unwin.

Bohn, O.-S. (2000). Linguistic relativity in speech perception: An overview of the influence of language experience on the perception of speech sounds from infancy to adulthood. In S. Niemeier & R. Dirven (Eds.), *Evidence for linguistic relativity*. Amsterdam: J. Benjamins, pp. 1-28.

Bohn, O.-S. & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11, 303-328.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131-158.

Borges, J. L. (1964). Funes, the memorious. In D. A. Yates and J. E. Irby (Eds.), *Labyrinths. Selected stories & other writings*. New York: New Directions, pp. 59-66.

Bowerman, M. (1988). The "no negative evidence" problem: How do children avoid constructing an overgeneral grammar? In J. A. Hawkins (Ed.), *Explaining language universals*. Oxford: Blackwell, pp. 73-101.

Broeder, P. & Murre, J. (Eds.) (2000). *Models of language acquisition: Inductive and deductive approaches.* Oxford: Oxford University Press.

Browman, C. P. & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155-180.

Brown, R. (1958). *Words and things*. Glencoe, Ill.: Free Press.

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, 10, 159-199.

Chomsky, N. (1964). Current issues in linguistic theory. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language*. Englewood Cliffs, NJ: Prentice-Hall, pp. 50-118.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Clark, E. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum, pp. 1-33.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303-306.

Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.

Fodor, J. A. (1980). The present status of the innateness controversy. In J. A. Fodor (Ed.), *Representations: Essays on the foundations of cognitive science*. Cambridge, MA: MIT Press, pp. 257-316.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Gillis, S., Daelemans, W., & Durieux, G. (2000). 'Lazy learning': Natural and machine learning of word stress. In P. Broeder & J. Murre (2000), *Models of language acquisition: Inductive and deductive approaches.* Oxford: Oxford University Press, pp. 6-99.

Gleitman, L. (1990). The structural sources of verb meaning. *Language Acquisition*, 1, 3-55

Goldsmith, J. (2001). Unsupervised learning of the morphology of a natural language. *Computational Linguistics*, 27, 153-198.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1166-1183.

Harnad, S. (2003). Categorical perception. In *Encyclopedia of cognitive science*. London: Nature Publishing Group/Macmillan. At http://www.ecs.soton.ac.uk/~harnad/Temp/catperc.html

Harris, R. (1973). *Synonymy and linguistic analysis*. Oxford: Blackwell.

Harris, Z. (1951). *Methods in structural linguistics*. Chicago: University of Chicago Press.

Hazan, V. & Barrett, S. (2000). The development of phonemic categorization in children aged 6-12. *Journal of Phonetic*s, 28, 377-396.

Heitner, R. M. (2005). An odd couple: Chomsky and Quine on reducing the phoneme. *Language Sciences*, 27, 1-30.

Hockett, C. (1942). A system of descriptive phonology. *Language*, 18, 3-21.

Hoijer, H. (1958). Native reaction as a criterion in linguistic analysis. *Proceedings of the Eighth International Congress of Linguistics*, 573–591.

Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In V. Fromkin (Ed.), *Tone: A linguistic survey*. Orlando: Academic Press, pp. 77-111.

Jaeger, J. J. (1980). Testing the psychological reality of phonemes. *Language and Speech*, 23, 233-253.

Jaeger, J. J. (1986). Concept formation as a tool for linguistic research. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental phonology*. Orlando: Academic Press, pp. 211-238.

Jaeger, J. J. & Ohala, J. J. (1984). On the structure of phonetic categories. *Berkeley Linguistics Society*, 10, 15-26.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Jones, D. (1929). Definition of a phoneme. *Le Maître Phonétique*, 43-44.

Jones, D. (1973). The history and meaning of the term "phoneme". In E. C. Fudge (Ed.), *Phonology*. Harmondsworth: Penguin, pp. 17-34.. First published 1957 in supplement to *Le maître phonétique.*

Joos, M. (Ed.). 1957. *Readings in linguistics*. Chicago: University of Chicago Press.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Kasabov, N. (2002). *Evolving connectionist systems: Methods and applications in bioinformatics, brain study and intelligent machines.* London, New York & Heidelberg: Springer.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69.

Kornai, A. (1998). Analytic models in phonology. In J. Durand & B. Laks (Eds.), *The organization of phonology: Constraints, levels and representations*. Oxford: OUP. pp. 395-418. Also at http://people.mokk.bme.hu/~kornai/Papers/roy1.pdf

Krámský, J. (1957). *The phoneme: Introduction to the history and theories of a concept*. Munich: Wilhelm Fink.

Kuhl, P. (1987). Perception of speech and sound in early infancy. In P. Salapatek & L. Cohen (eds.), *Handbook of infant perception*, Vol. 2. New York: Academic Press, pp. 273-382.

Kuhl, P. (1991). Human adults and human children show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50, 93-107.

Lachs, L., McMichael, K., & Pisoni, D. B. (2000). Speech perception and implicit memory: Evidence for detailed episodic encoding of phonetic events. *Progress Report 24. Speech Research Laboratory*, Indiana University, Bloomington, pp. 149-167.

Langacker, R. W. (1987). *Foundations of cognitive grammar,* Vol. 1*: Theoretical prerequisites*. Stanford: Stanford University Press.

Lehiste, I. & Peterson, G. E. (1961). Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America*, 31, 428-435.

Liberman, A. M., Cooper, F. S, Shankweiler, D. P. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Liberman, A. M., Delattre, P. D., & Cooper, F. S. (1958). Some cues for the distinction between voiced and unvoiced stops in initial position. *Language and Speech*, 1, 153-167.

Liberman, A. M. & Mattingly, I. G. (1985). The motor theory of speech perception revisited. *Cognition*, 21, 1-36.

Lindblom, B. (2000). Developmental origins of adult phonology: The interplay between phonetic emergents and the evolutionary adaptations of sound patterns. *Phonetica*, 57, 297-314.

Lisker, L. (1978). In qualified defense of VOT. *Language and Speech*, 21, 373-383.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops. *Word*, 20, 384-422.

Lotto, A. J. (2000). Language acquisition as complex category formation. *Phonetica*, 57, 189-196.

Luria, A. R. (1968). *The mind of a mnemonist*. Cambridge, MA: Harvard University Press

McLeod, P. Plunkett, K. & Rolls, E. T. (1998). *Introduction to connectionist modelling of cognitive processes.* Oxford: Oxford University Press.

Manning, C. & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Mattingly, I. (1972). Reading, the linguistic process, and linguistic awareness. In J. Kavanagh & I. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading*. Cambridge, MA: MIT Press, pp. 133-147.

Maye, J., & Gerken, L. (2000). Learning phonemes without minimal pairs. *Proceedings of the 24th Boston University Conference on Language Development*, 522-533

Maye, J., Werker, J. F. & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101-B111.

Mompeán, J. A. (2002). *The categorisation of the sounds of English: Experimental evidence in phonology*. Unpublished Ph.D. dissertation, University of Murcia.

Mompeán, J. A. (2004). Category overlap and neutralization: The importance of speakers' classifications in phonology. *Cognitive Linguistics*, 15, 429-469

Morales, F. & Taylor, J. 2005. Learning and relative frequency. Unpublished manuscript, University of Otago.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Murphy. G. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Murphy, G. & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316.

Nathan, G. (1986). Phonemes as mental categories. *Berkeley Linguistics Society*, 12, 212-223.

Pater, J., Stager, C., & Werker, J. (2004). The perceptual acquisition of phonological contrasts. *Language*, 80, 384-402.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24, 175-184.

Peterson, G. & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, 32, 693-703.

Pierrehumbert, J. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speec*h, 46, 115-154

Pike, K. (1947). Grammatical prerequisites to phonemic analysis. *Word*, 3, 155-172.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Plunkett, K. (1995). Connectionist approaches to language acquisition. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language*. Oxford: Blackwell, pp. 36-72.

Postal, P. (1968). *Aspects of phonological theory*. New York: Harper & Row.

Potter, R. K., Kopp, G. A. & Green, H. C. (1947). *Visible speech*. New York: Van Nostrand.

Quine, W. v. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Quine, W. v. O. (1987). *Quiddities: An intermittently philosophical dictionary*. Cambridge, MA: Harvard University Press.

Rumelhart, D. & McClelland, J. (1986). On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In J. McClelland, D. Rumelhart, and the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. 2. Cambridge, MA: MIT Press, pp. 216-271

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.

Sampson, G. (1985). *Writing systems: A linguistic introduction*. Stanford, CA: Stanford University Press.

Sapir, E. (1921). *Language: An introduction to the study of speech*. New York : Harcourt, Brace & World.

Saussure, F. de (1964). *Cours de linguistique générale*. Paris: Payot. First published 1916. Translated by R. Harris, as *Course in General Linguistics*. London: Duckworth (1983)

Schane, S. (1971). The phoneme revisited. *Language*, 47, 503-521.

Shillcock, R., Cairns, P., Chater, N., & Levy, J. (2000). Statistical and connectionist modeling of the development of speech segmentation. In P. Broeder & J. Murre (Eds.), *Models of language acquisition: Inductive and deductive approaches.* Oxford: Oxford University Press, pp. 103-120.

Slobin D. I. (1985). Cross-linguistic evidence for the language-making capacity. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition.* Vol 2*: Theoretical issues*. Hillsdale, NJ: Erlbaum, pp. 1157-1249.

Taylor, J. R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.

Taylor, J. R. (2003a). *Linguistic categorization*. Oxford: Oxford University Press. First edition: 1989.

Taylor, J. R. (2003b). Near synonyms as co-extensive categories: 'high' and 'tall' revisited. *Language Sciences*, 25, 263-284.

Taylor, J. R. (to appear). Prototypes in cognitive linguistics. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*. Mahwah NJ: Lawrence Erlbaum.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge MA: Harvard University Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Treiman, R. (1985). Phonemic awareness and spelling: Children's judgments do not always agree with adults. *Journal of Experimental Child Psychology*, 39, 182-201.

Treiman, R. & Baron, J. (1981). Segmental analysis ability: Development and relation to reading ability. In G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice*, Vol. 3. New York: Academic Press, pp. 159-198.

Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Oxford: Blackwell.

Weitzman, R. (1992). Vowel categorization and the critical band. *Language and Speech*, 35, 115-125.

Werker, J. (2003). The acquisition of language specific phonetic categories in infancy. *Proceedings of the 15th International Congress of Phonetic Sciences*, 21-26.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49-63.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Phonological Concepts and Concept Formation: Metatheory, Theory and Application[1]

HELEN FRASER
University of New England[*]

ABSTRACT

This paper presents an overview of Phenomenological Phonology (PP), including its metatheory, theory and application, for comparison with Cognitive Phonology (CP). While PP and CP are in close agreement at the theory level, there are some significant differences at the level of metatheory. PP considers phonological terms (such as *phoneme* and *word*) to be words like any others, and gives detailed consideration to the concepts behind such terms. It also considers pronunciation to be a form of behaviour, driven by concepts created through general concept-formation processes. This has important consequences for practical application in the areas of pronunciation and literacy teaching.

KEYWORDS: phenomenology, phonology, pronunciation, concepts, concept formation, abstractness, applied cognitive phonology.

[*] ***Address for correspondence***: Helen Fraser. School of Languages, Cultures and Linguistics. University of New England, Armidale, New South Wales, Australia. Tel: 61 2 6773 3318; e-mail: hfraser@une.edu.au; web: http://www-personal.une.edu.au/~hfraser/

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I. INTRODUCTION

This paper offers an overview of Phenomenological Phonology, with the aim of inviting comparison with Cognitive Phonology (Langacker, 1987; Taylor, 2002). Phenomenological Phonology (PP) is similar in many ways to Cognitive Phonology (CP), but there are also a number of differences. Should PP be thought of as the same as CP, a branch of CP, or something different altogether? And does anything much hinge on the answer?

Since an important feature of PP is its explicit linking of metatheory, theory and practice, this paper, despite space limitations, covers key aspects of all three. The first part gives an overview of the metatheoretical framework (for more detailed background see Fraser, 1992, in press). The second part briefly presents some implications of the framework for phonological theory (see also Fraser, 1997, 2004b). Finally, some examples are given of the effect of following these implications on the practical task of pronunciation teaching (Fraser, 2001, 2004a, 2006).

II. THE PHENOMENOLOGICAL FRAMEWORK

II.1. Overview

Phenomenology is a complex set of philosophical ideas (Spiegelberg, 1982). This paper focuses on just one general aspect, its recognition of three distinct levels of analysis: the level of words, the level of reality and, mediating between the two, the level of concepts.

These three levels of analysis are very familiar to Cognitive Linguistics, partly because of the influence of phenomenology, through structuralism and post structuralism, on contemporary scholarship in general, and partly because linguistics itself, through the work of key figures such as Saussure, Whorf and their successors, has played an important role in developing a detailed understanding of the three levels and their relationships (Carroll, 1956; de Saussure, 1986). However, phenomenology pursues implications of the three levels of analysis that are not so familiar.

The most important of these implications follows from recognition that words and concepts do not emerge from nothing, but are created and used by a person. Recognition of the levels therefore requires recognition of a Subject. The Subject is a generalised and theorised version of the person who creates and uses words and concepts, as opposed to a subject (lower case ‘s’), which is a specific person. There is thus an important distinction between the terms *Subjective*, meaning ‘requiring theoretical acknowledgement of a

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

'someone' who creates concepts and words', and *subjective*, meaning 'restricted to the viewpoint of a specific person'.

The Subject is acknowledged by a range of theories, but Phenomenology is distinctive in putting the Subject in central position, starting any analysis with explicit consideration of the Subject who creates the words and concepts used in the analysis. This can be especially useful when, as in the scientific study of language and cognition, there are two Subjects: one, the scientist, making a theoretical analysis of the other, an element in the scientist's theory.

A less well known but very significant contribution of Phenomenology, stemming directly from its focus on the role of the Subject in creating words and concepts, is observation of the ease with which the three levels of analysis (word, concept and reality) can become confused, and the undesirable effects on theory that can follow. More than simply observing this, Phenomenology has developed a method for avoiding this confusion.

The following sections provide more detail about each of the three levels of analysis, the role of the Subject, and the method for minimising confusion among the levels. Some of the ideas will be familiar to readers with a background in Cognitive Linguistics. However the use made of the ideas in PP is, as will be seen, somewhat different.

II.2. Reality

The level of reality is covered by a range of technical terms in Phenomenology (e.g. *World*, *Life World*), each with important distinctions in meaning, but the everyday word *reality* serves well for the present discussion. Reality is the world as it is, as opposed to how people think it is or would like it to be, or how it might potentially or ideally be. People and their artefacts are part of reality and can influence reality, but reality exists independently of any individual. Phenomenology is therefore not a form of idealism (the assertion that reality exists only in people's minds).

Reality is richly structured and highly complex, with a nature, or way of being, of its own. Through embodied experience, people can develop an unspoken or tacit understanding of reality. This understanding can be considered accurate, or realistic, to the extent that it allows people to survive and prosper as individuals and as a species. It is, however, inevitably partial, both in the sense of being limited, and in the sense of being subjective, constrained by the sensory systems and interpretive biases of the experiencer.

It is also possible to describe reality with words or other symbols and representations. This is useful in communicating with others about reality but, since descriptions rest

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

ultimately on experience, they too are inevitably partial, again in both senses of the word. Reality is open to many alternative descriptions. Through the use of special methods and equipment, scientists can create descriptions of greater and greater detail and accuracy. However, reality is inexhaustible by any, or all, descriptions. There is always some 'surplus' beyond the description, and any one description inevitably obscures some aspects of reality, even as it may be revealing others.

These points can be summed up by saying that reality in itself is 'raw'. It has a nature of its own, but no self-existing description of its own. Ultimately, even the most detailed and accurate descriptions of reality rest on a tacit appeal to shared experience.

II.3. Concepts

The level of concepts is also given various technical terms in Phenomenology, as well as in other theories. Perhaps the best known term is *mental representation*. However, this term has a good deal of theoretical baggage (Shanon, 1993). Though the word *concept* was disallowed in technical discussion for some time due to the unobservability of concepts, it has now made a welcome return, partly thanks to Cognitive Linguistics (Taylor, 2002).

Concepts are the Subject's interpretation of, or way of thinking about, reality. Everyday reasoning about people's thinking and behaviour involves constant reference to concepts in understanding why people do what they do, and predicting what they might be likely to do next, via an informal 'theory of mind' (Premack & Woodruff, 1978).

One of the most important things about concepts is that it is concepts of reality, not reality itself, that drive a person's behaviour. People sometimes find this difficult to accept, preferring to explain behaviour, especially their own, with reference to reality ('I ran away because of the tiger') rather than with reference to their concepts ('I ran away because of my concept of the tiger'). It is not too difficult to demonstrate, however, that it is not the tiger as such which causes the running, but a concept of the tiger (the tiger would not cause that behaviour in someone who did not know it was there, or did not know it was dangerous).

Though concepts have an immense effect on reasoning and behaviour, concepts themselves are not directly observable. Indeed there is a sense in which they are necessarily invisible. Concepts are like a lens, or a pair of glasses, through which a person views the world. They greatly affect the person's view of reality, but when people use concepts, they look through them rather than at them, and are generally unaware of them. When glasses are not in use, they can be taken off and looked at. Unfortunately concepts cannot be directly

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

observed in this way. They can be studied through consideration of their effect. However, as with glasses, it is difficult to study concepts while actually using them.

Given their central importance to all aspects of life, concepts as such have been given surprisingly little direct study in the mainstream sciences of cognition, which have followed an analogy between human cognition and computer processing, and sought to avoid reference to concepts in favour of mental representations, which are, even if only principle, more tangible. Cognitive Linguistics is exceptional in taking conceptual structure as a central theme in theory development, and has contributed to a body of scientific knowledge about concepts, which includes the following general points (Murphy, 2002).

Concepts are abstract with respect to reality. This is because they are formed through processes of abstraction from reality. Abstraction is a 'drawing out' (the etymological meaning of *abstract*) of those aspects of reality that are salient to a particular person at a particular time in a particular context. Abstraction requires a sense of contrast, or difference. It is common to say that a concept groups together aspects of reality that are similar. It is equally true, however, that concepts group together aspects of reality that are different from some other known aspect of reality. For example, developing the concept BROWN involves understanding what is 'not-brown' (Wittgenstein, 1958/1974). Cognitive Linguistics, following the work of Eleanor Rosch (e.g. Rosch, 1973), defines concepts in terms of categories, with 'prototypical' members at the centre of the category and more 'peripheral' members around an often 'fuzzy' boundary.

The fact that the creation of concepts involves processes of abstraction means, importantly, that concepts are not simply a 'mapping' of physical characteristics of reality onto another level. Concepts are strongly influenced by context, culture and point of view - as well as by reality itself. The latter is important to emphasise because Phenomenology has often been wrongly thought of as focusing on subjective (note the small *s*) interpretation to the exclusion of reality ('if I think it is art, it is art'). The phenomenological understanding of the relationship between concepts and reality can be seen via analogy with the potter and the clay. Reality is the clay, and the concept is the pot. The potter can shape the clay in many ways, but the nature of the pot is constrained by the nature of the clay. Similarly, reality can be conceptualised in many ways, but the nature of the resulting concepts is constrained (barring pathology) by the nature of reality.

The fact that concepts do not simply map aspects of reality means that, while it can be useful to work with formal definitions of concepts for certain purposes, no concept can be fully defined with reference only to the physical or formal properties of the reality behind it

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

(Wittgenstein, 1958/1974). Rather, a concept embraces a collection of aspects of reality that seem 'the same' to a particular subject in a particular context, despite having often enormous physical differences. As is well known in Cognitive Linguistics, it is quite possible to have a category in which there is no single physical feature common to all members.

II.4. Words

The level of words is one with which everyone is familiar from everyday life. When people use words, they feel that they are referring directly to reality. However, this is an illusion, sometimes known as *naive realism*. Words refer not to reality but to concepts of reality.

Words, like pictures, diagrams and other representations, are symbols not of reality but of concepts. The ability to use such symbols is an important part of what makes us human (Noble & Davidson, 1996). Words have the important function of allowing humans to reflect upon concepts, and think about the things they conceptualise even when they are not there. With the aid of symbols, people can compare and contrast concepts, consider similarities and differences among them, and draw out, or abstract, those features that are salient at the time and in the context. This gives the powerful ability to iteratively create new concepts at higher levels of abstraction. For example, having words for concepts of things that happen to be red and blue allows comparison of the things and abstraction of new, more abstract, concepts of RED and BLUE. By further comparing and contrasting, yet more abstract concepts such as COLOUR, or SCARLET, CRIMSON, AQUA, and ROYAL BLUE, can be created.

By considering the processes of abstraction involved in the formation of concepts behind words, a hierarchy of abstractness can be defined. This hierarchy is similar to but somewhat different from the schema-instance hierarchy used in Cognitive Linguistics (Taylor, 2002: Chapter 7). Basic concepts are similar to the basic concepts of Cognitive Linguistics, but concepts of both lower order and higher order terms in a taxonomy are considered in PP to be more abstract than basic level concepts, because they require more levels of abstraction by the Subject. Thus for example, both TOOL and CHAINSAW would be considered more abstract in PP than SAW.

Applying a word to a concept, like creating a concept in the first place, is a process that requires tacit background understanding (Polanyi, 1966). Because of this, just as concepts are not a direct mapping of reality, so words are not a direct mapping of concepts. One concept can be symbolised by several words (synonymy), or one word can symbolise several concepts

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

(homonymy). It is possible to have a concept without having a word for it. Some concepts require not a word, but a whole sentence, for their expression.

This is the reason behind the observation, very well known to Cognitive Linguistics, that determining the concept behind a word (its meaning) is not simply a matter of reading off a formal definition, but requires analysis of its use in a range of contexts.

II.5. The Subject

From what has been said so far, it is clear that the existence of concepts and words presupposes some processing of reality, in much the same way as the existence of a pot presupposes some processing of clay. The Subject is the doer of the processes of conceptualisation, abstraction, categorisation, and so on, conceived as a generalised theoretical being, rather than a specific person. It is important to clarify this definition. It is very far from implying that subjects (people) have only general or group characteristics, and not individual, personal characteristics. It was the major contribution of Heidegger to Phenomenology, in opposition to Husserl, to point out that personal, social and embodied characteristics were essential to people's interaction with reality, and their ability to form concepts and use language. The general characteristics of the Subject therefore include provision for highly specific characteristics of each subject, based on its own embodied, social and subjective experience.

Without the Subject, clearly, there could be no concepts or words. Interestingly, however, just as concepts are necessarily invisible as they are being used, so the Subject is invisible to itself while engaged in its projects. It takes an act of reflection for a Subject to become aware of its own characteristics, and the contribution of its own point of view to its concepts of reality. Even then, such awareness is necessarily partial.

One of the Phenomenologists' major contributions is their acknowledgement of and focus on the role of the Subject in creating words and concepts. This is not to say that the role of the Subject is denied in other philosophies. However, in other philosophies, the Subject, even if acknowledged in principle, tends to be regretted and avoided. A great deal of effort goes into defining the Subject so as to minimise any difference between the Subject and the material world (Stillings *et al*., 1987). Mainstream cognitive theory, for example, has achieved this avoidance of dualism via analogy between the human mind and a computer.

Phenomenology, in strong contrast with mainstream theory, does not seek to avoid the Subject, but embraces the Subject, acknowledging the fundamental difference between the

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Subject and other aspects of reality, and analysing with care the characteristics of the Subject that allow it to achieve processes such as conceptualisation, abstraction or categorisation. One important finding is that, since the processes involved in formation and use of concepts and words require a tacit background understanding not just of the concept itself but of the context in which it emerges (Polanyi, 1966), the Subject who creates words and concepts cannot be a computational device (Dreyfus, 1979).

Through its recognition of the Subject, and analysis of the relationship between reality, concept and Subject, Phenomenology is able to transcend the traditional philosophical opposition between idealism and realism, while still avoiding dualism. It does this by distinguishing between *existence* and *existence-as*: an aspect of reality can *exist* in a 'raw' or undescribed state independently of any observer or Subject, but for it to *exist-as* (some description) - that is, to have a word attached to it - requires action, in particular the action of concept-formation, by a Subject.

Most importantly of all, phenomenologists see themselves, in their role as philosophers, as Subjects in exactly the same sense as the Subject they postulate in their philosophy, and the terms of their theories as words in exactly the same sense as words of everyday language.

II.6. The natural attitude

One of Phenomenology's most useful contributions, though one that has influenced other disciplines relatively little, is their concept of the Natural Attitude (Husserl, 1960). The Natural Attitude is the tendency to behave in everyday life as if words refer not to concepts but to reality.

The Natural Attitude is somewhat like naive realism. However, it is not a theoretical 'ism' but an informal attitude. As its name suggests, it is entirely natural, an inevitable consequence of the invisibility of concepts and the tendency of the Subject to focus outwards from itself when engaged in projects. People actually need the Natural Attitude when getting on with their lives and engaging in projects. There are many times, for example, when it is much better to accept the convenient fiction that the word *dog* refers to an actual dog than to waste time on reminders that it really refers to someone's concept of a dog.

On the other hand, it is possible to choose a different attitude at times. In the Attitude of Reflection, the fictions of the Natural Attitude can be put aside, in an attempt to see the effect of one's concepts on one's view of the world, in much the same way as one can consider the effect of one's glasses on one's visual perception.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

In everyday life, people move without concern between the Natural Attitude and an attitude of reflection. For example, the everyday terms *sunrise* and *sunset*, with their Natural Attitude view of the sun revolving round the earth, continue to be used even though scientific discourse requires acceptance of the reflective attitude view that it is the earth that revolves around the sun.

However, though in fairly neutral contexts like this, the move between Natural and reflective attitudes is easy, there can sometimes be resistance to the reflective attitude view. For example, it is one thing to accept the general principle that behaviour is driven by concepts, not reality, or to say it of someone else's behaviour. It is another to admit that one's own behaviour is driven by concepts, not reality. There is a sense in which people cling to, or get stuck in, the Natural Attitude.

Unfortunately scientists and philosophers are far from immune from getting stuck in this way. They can recognise the general principle that words relate to concepts not reality, and yet have a wish to behave as if their own scientific words relate directly to reality, justifying the use of these words by reference to their accurate portrayal of reality. This can cause particular irony when scientists or philosophers studying words and concepts use words and concepts in a way that conflicts with their own findings about words and concepts - which is why the Phenomenologists take such care to recognise that they themselves, as theorists, are Subjects just like the Subjects they study.

II.7. Theorising words and concepts

Everyday reasoning, as we have seen, involves an understanding of words and concepts that is very similar to the one just outlined, and makes frequent reference to the role of words and concepts in reasoning and behaviour. It is common knowledge that different people can have different concepts of the same 'raw reality', depending on their culture, point of view and context. If someone says 'It's a great movie', we do not accept that as an objective statement of fact, but interpret it in light of who has said it, and in what context, and come to our own judgment about the likely quality of the movie. In contemporary debate about important social or political issues, the role of language and context in shaping opinion is extremely well understood - perhaps at least partly through the influence of linguistics (Elgin, 1999).

Theories of language and cognition seek more scientific understanding of the role of words and concepts. Mainstream theories have a rather specific understanding of what it means to be scientific, based on the practices of the natural sciences. The natural sciences,

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

recognising that words and concepts intervene between reality and people's understanding of it, have traditionally sought to bypass words and concepts in attaining an understanding of reality itself. This is achieved through ensuring that all statements are based on verifiable observations. Of course it is recognised that the ability to achieve complete objectivity is limited (Feynman, 1986). However the ongoing attempt is seen as the path to scientific understanding of the natural world.

Applying this criterion to the scientific study of language and cognition however creates some problems. It is of course possible and necessary to treat language and cognition in a scientific manner, but in doing so it is not possible to bypass words and concepts, since they are key elements of the subject matter being studied, as well as the vehicle via which they can be studied. Traditionally rigour in theories of language and cognition has come from explicit definition of terms, and justification of the definition by appeal to the relationship between the terms and the reality they refer to.

The problem is, of course, that according to these theorists' own understanding of language and cognition, there is no direct relationship between reality and words. The relationship must always be mediated through concepts - which must always belong to an embodied Subject in a personal and social context. The natural sciences can get away (up to a point) with ignoring this uncomfortable fact, by behaving as if all scientific concepts belong to some generalised 'scientific Subject' with a particular point of view that all scientists are willing to subscribe to. In the sciences of language and cognition, however, there are unavoidably two Subjects to consider, the one being studied and the one doing the studying. Terms within the theory need to be defined from the point of view of one or the other, and the issue cannot be fudged for long without theoretical confusion arising.

For this reason, the phenomenological approach to rigour is different. The traditional approach is an attempt to escape from the Natural Attitude by denying it. According to Phenomenology, the only hope of avoiding the bad effects of the Natural Attitude on theory is to acknowledge it. Rigour in Phenomenology, then, comes from careful analysis of the words and concepts used in theories, not to eliminate their Subjectivity but to understand their presuppositions and ensure that these are commensurate with the context in which they are being used.

The phenomenological method therefore involves not defining the terms used in theories once and for all, but asking Framework Questions of each term in the context that it is used. Framework Questions are questions like the following:

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

- what concept lies behind this term?
- what 'bit of reality' lies behind this concept?
- what kind of person can have this concept?
- what prior concepts does that person need to have?
- in what context does that 'bit of reality' have to occur to be conceptualised in this way?

Through asking questions like this, a better understanding can be gained of the relative abstractness and other characteristics of the concepts behind the words used in theories.

III. PHENOMENOLOGICAL PHONOLOGY

III.1. Overview

Phonology is the study of the sounds of speech, and how they function to help us convey meaning in language. It has been barely touched as a topic of study in Phenomenology (Ihde, 1976; Merleau-Ponty, 1962), perhaps because most of the interesting questions of phonology arise from technical discoveries (Perkell & Klatt, 1986) that have impinged little on the more philosophical work of the Phenomenologists.

Mainstream phonology, on the other hand, has had almost no exposure, even indirectly, to Phenomenology. Even those aspects of phenomenological thinking that are well known in other branches of linguistics (e.g. the idea that words relate to concepts not to reality, and that concepts cannot be defined simply by listing physical features of the reality they relate to) are rarely the focus of attention in phonology. Thus it is absolutely standard in mainstream phonology to define phonemes as a set (whether a list or a 'geometry') of physical features, and to understand phonetic features as being 'closer to reality' than phonological features. Cognitive Phonology, again, is exceptional in the degree to which it has questioned these mainstream ideas, with several scholars looking at the implications of treating phonemes not as sets of features but as categories of sounds (Nathan, 1986, 1996; Taylor, 2003).

This section summarises very briefly an investigation of how phonology might be treated from a phenomenological perspective.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

III.2. Phonological terms

The first step in developing phonology from a phenomenological perspective is to acknowledge that the terms used in a theory of phonology (terms such as *word*, *syllable*, *phoneme*, *feature*, *alveolar*, and so on) are words like any other words, subject to all the principles discussed above. They refer to concepts of reality, not to reality itself, and people who use them are prone to get caught up in the Natural Attitude. It is essential therefore to ask the Framework Questions about the terms used in phonological theories.

The reality behind any phonological terms is what might be called 'raw speech' -- the rich, complex, highly structured, quasi-continuous sound produced by the vocal tract when someone talks. Of course, to give a description of any kind to raw speech requires first conceptualising it, thus rendering it no longer 'raw', and limiting it to only one of its possible descriptions rather than any of the many others. One way to refer to raw speech without limiting it in this way, is to 'point' to the shared experience of hearing speech without understanding the words, for example when listening to someone speak an unknown language. The aim of the exercise of course is not describe the raw speech neutrally (which is impossible), but rather to demonstrate the degree to which everyday descriptions of speech in terms of words are abstract, or 'processed', with respect to the raw reality - in the senses developed above.

An analogy with the crow of the rooster is sometimes helpful in elaborating the distinction between raw speech and words. The rooster can be described as saying 'cockadoodle-doo' (or its equivalent in another language). It can also be experienced as uttering a 'raw crow' - a very different sound from the English word 'cockadoodle-doo'. Similarly, a person can be described as saying 'Could you pass the salt please?' In this case it is more difficult to discern a difference between the words and the reality, but it is possible to think of the 'raw speech' behind the words by imagining how the sentence might seem to someone who does not know English, or to an animal which does not understand language, or to a machine set up to record the speech.

Between raw speech and the phonological terms used to describe it are, of course, concepts - as is the case for any words. There are many different ways of conceptualising raw speech. There is nothing inherently right or wrong about any of these concepts, but each has presuppositions that need to be taken into account when using them in theories.

Before raw speech can be understood as meaningful language, it must first be conceptualised as a string of words. It is interesting that in everyday life, people rarely

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

actually describe speech as a sequence of words. Conceptualising speech as a sequence of words has become so obvious to them, through learning a native language, that they no longer notice they do it, and have few metalinguistic terms to describe the sounds of words, as opposed to their meanings.

One common way for phonologists to conceptualise speech is as a string of phonemes. Asking the framework questions shows that in order to have a concept of phoneme, it is necessary to know a large number of words, and to be able to compare and contrast the sounds of those words in a very particular manner (Byrne, 1998), which may only become possible through learning alphabetic literacy (Coulmas, 1989; Linell, 1988; Olson, 1996). The concept of phoneme thus presupposes a Subject that knows the language, is literate in an alphabetic writing system, and, arguably, has undergone some basic training in linguistics (Scarborough *et al*., 1998).

Another common way for phonologists to conceptualise speech is as a sequence of allophones. Pursuing the framework questions reveals that the concept of allophone presupposes a concept of phoneme. Raw speech, as discussed earlier, is continuous, and continuously variable. It is not possible to segment it into allophones without having some prior understanding of its phonemic segmentation (Laver, 1994).

Although these and other conceptualisations of speech seem very obvious to phonologists, it is not difficult to demonstrate that the most basic way for someone who knows the language to conceptualise speech is as neither phonemes nor allophones, but as a sequence of meaningful words. This is shown by the fact that it is almost impossible to listen to a language you know without hearing it as words. Though these word-concepts are not noticed, as discussed above, recognition of them underpins all other concepts of the sound of speech, which are more abstract.

III.3. Abstractness of phonological terms

Using answers to the Framework Questions along the lines just indicated, it is possible to set up a hierarchy of abstractness of phonological terms, based on understanding of the processes of abstraction required to create the concepts the terms refer to.

From all that has been said so far (and see also more detail in the references given in the Introduction) it is clear that the metalinguistic concept behind the term *word* is abstract with respect to raw speech. Terms for parts of words, such as *phoneme* or *syllable*, refer to

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

considerably more abstract concepts, and terms for phonetic concepts such as *allophone* or *pitch trace*, are more abstract still.

Compelling as this view is when argued from the principles outlined above, however, it is highly unorthodox in relation to mainstream phonological theory. Mainstream theory would certainly agree that words are abstract with respect to raw speech. However it would see phonetic representations as less rather than more abstract than words, closer to the reality of actual pronunciation. Phonemic representation is universally agreed, among mainstream theories, to be more abstract than phonetic, and words more abstract again, because of their even less direct mapping onto raw speech.

How are these two strongly opposed views to be reconciled? It is natural to think that one must be right and the other wrong. However this may not be the most helpful approach. Rather than seeking universal accuracy or objectivity, it is possible to recognise that there are various ways of conceptualising abstractness, each with a range of presuppositions.

In many contexts the traditional hierarchy is perfectly appropriate. For example, if linguists are theorising language for descriptive or typological purposes, intending to communicate findings primarily with other phonologists who have a common goal of predicting and accounting for characteristics of the sound systems of language, it is very useful to be able to refer to rules that change one sound into another, or constraints that restrict the appearance of certain sound combinations -- without worrying whether sounds really change into other sounds, or whether it is people who change the way they speak, or become aware of different relationships among sounds (Ohala, 1990).

In other contexts, rather than acting merely as Subjects theorising language, linguists act as Subjects theorising other Subjects' use of language. Sometimes this is done in a context, such as in developing phonological theory for speech technology, which makes an analogy between the Subject and a computational device. In this case as well the traditional hierarchy is appropriate to the extent it is useful in achieving the goal.

A crucial change of context comes, however, when linguists aim to theorise a living human being seeking to accomplish some phonological task, such as acquiring literacy or learning pronunciation. To model such a Subject as a computational device is to distort its nature. To maintain the mainstream hierarchy of abstractness is to ascribe concepts to Subjects at stages of development, or stages of mental processing, at which it is impossible that such a concept could exist.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The important thing is to choose a hierarchy of abstractness, and other theoretical commitments, whose presuppositions are commensurate with the context in which it is to be used. Asking the Framework Questions can help with this.

IV. APPLYING PHENOMENOLOGICAL PHONOLOGY TO PRONUNCIATION TEACHING

It is well known that people learning a second language in adulthood often have particular difficulty mastering its pronunciation. It is often taken for granted that this difficulty stems from the physical difficulty of producing the sounds of the new language. On this view, teaching pronunciation involves showing students 'the sounds you are actually making', via acoustic or articulatory representations. This, unfortunately is often far less successful than is hoped. From the PP point of view, that is hardly surprising. Although in phonology articulatory and acoustic descriptions of speech are seen as being 'close to reality', according to the PP hierarchy, as discussed earlier, these are highly abstract representations, and a great deal of prior knowledge is required in order to interpret them, and to associate them with particular ways of speaking. Learners of second language pronunciation typically do not have that prior knowledge, and find it very difficult to relate the visual representations to their own pronunciation behaviour.

PP takes a different view of how to teach pronunciation. Pronunciation is a form of behaviour. As such, it is driven by concepts. On this view, difficulties with pronunciation are primarily conceptual difficulties. This is supported by the observation that, though of course some pronunciation problems are caused by physical difficulty in producing particular sounds, in many cases, the speaker has no difficulty producing an acceptable version of the needed sounds. A classic case is the speaker who calls two girls *Arison* and *Blonwyn*. This person can clearly say both [r] and [l]. The problem is not in pronouncing the sounds, but in keeping them mentally distinct, as appropriate for the new language.

If pronunciation is driven by concepts, the key to changing pronunciation is changing the concepts that drive it. This will not instantly solve all pronunciation difficulties. Pronunciation is a skill, and requires practice. However, without attention to the conceptual level, practice alone is frequently ineffective, and therefore discouraging. If the conceptual issues are addressed first, practice can be rewarding, and improvement follows far more quickly (Couper, 2006).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Principles for helping students form or modify concepts are well known to education (Lefrancois, 1994), and especially to communicative language teaching (McKay, 2002). They include principles such as ensuring that materials are meaningful, contextualised and culturally appropriate, that students are actively involved in learning rather than passively receiving information, that learning follows a series of incremental steps building new concepts from existing concepts, that students take responsibility for their own learning through self-reflection and other meta-cognitive skills. Such principles fit extremely well with PP and can be easily adapted to pronunciation teaching, by someone who understands principles of PP. Unfortunately however the major shift to communicative language teaching that took place in the 1970s and 1980s did not give detailed consideration to pronunciation (Celce-Murcia *et al*., 1996). This may have been because at that time there was a gulf between theoretical phonology, which was strongly influenced by the computational analogy, and language teaching, which required a more humanistic perspective. The result is that, while many teachers successfully use methods of teaching pronunciation that conform to a greater or lesser degree to PP principles (Fraser, 2000), many do not consider themselves successful pronunciation teachers, and many do not teach pronunciation at all (Macdonald, 2002).

Even when concept formation practices are used successfully, however, they are generally not understood explicitly as concept formation practices. This is because, when second language pronunciation teaching is theorised, it is still generally from the perspective of mainstream phonological theory. Applying mainstream computational theory in the classroom emphasises the need to teach phonological rules. Teaching such rules has its place, and can be successful if used sensitively. However it is all too easy for students to focus on learning the rules as abstract facts, rather than learning pronunciation. The result can often be that students can recite rules of pronunciation in pronunciation that itself violates the rules.

PP allows theorists to start with observation of what works best in the classroom, and create explanations which in turn allow the successful practice to be explored and extended in useful ways. The remainder of this section looks very briefly at just two principles of PP that can be used effectively in pronunciation teaching.

The first has to do with metalinguistic communication -- the communication that takes place between teacher and students about pronunciation. It is an obvious principle of concept formation that students should be able to understand what their teacher tells them about the concepts they are learning. It can be surprising to find, then, just how frequently students' and teachers' descriptions of pronunciation pass each other by completely, and how challenging it can be to remedy this and ensure successful metalinguistic communication.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Thus to tell a student that they have put the stress on the wrong syllable may seem to be a clear statement of fact. However people from different language and education backgrounds may have a very different concept of what phonetic characteristics constitute stress, or no concept of stress at all. Unless the teacher has taken care to build up the concepts that the student needs to share, it can prove to be very difficult for the student to understand and act on the teacher's advice.

People from different language and literacy backgrounds have very different phonological concepts (Strange, 1995). Communicating about pronunciation is therefore fraught with many opportunities for misunderstanding. It can be useful to think of metalinguistic communication as a form of intercultural communication. Metalinguistic communication is even more challenging than general intercultural communication, however. People tend to be aware of the possibility of different cultural concepts behind words like 'polite', or 'respect', but descriptions of sounds are thought to be objective, and misunderstanding is not expected.

A second useful principle brought into focus by attention to principles of concept formation is the use of contrast. As discussed above, contrast is essential for concept formation - but what sort of contrast is most effective for learners? Minimal pairs have been used in pronunciation teaching for many years (Baker, 1981), and can certainly be useful if incorporated into meaningful contexts. However minimal pairs are a very specific and rather abstract form of contrast. Another way to exploit the principle of contrast in pronunciation teaching is to focus on the contrast between what the learner thinks they said and what a native speaker would think they had said (Cartwright & Fraser, forthcoming).

V. CONCLUSION

Having started with a question as to the relationship between Cognitive Phonology and Phenomenological Phonology, it may be appropriate to finish with a suggestion as to how the question might be answered.

It seems the two approaches are very similar. CP is more established as a theory, within the broader framework of Cognitive Linguistics. Indeed PP has benefited greatly from the insights of CP, as well as Cognitive Linguistics more generally. PP perhaps has the advantage of its origins as a re-creation from first principles of the fundamentals of phonology in light of insights from both Phenomenological philosophy and pronunciation teaching practice. This has enabled PP to question and in some cases reject assumptions from mainstream phonology

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

which CP, despite its differences from mainstream theory, still accepts as axiomatic. The relative abstractness of concepts such as WORD, PHONEME and ALLOPHONE is a key example. More detailed discussion of this and other issues, along with suggested implications for applying Cognitive Phonology in the socially relevant domains of pronunciation and literacy teaching practice, are set out in further work by the author (Fraser, submitted).

NOTES

1. This paper is based on a presentation given at the Phonology Theme Session, International Cognitive Linguistics Conference, Seoul in July 2005. The author thanks the audience, as well as anonymous reviewers of this paper, for useful comments and suggestions.

REFERENCES

Baker, A. (1981). *Ship or sheep?: An intermediate pronunciation cours*e (2nd edition). Cambridge: Cambridge University Press.

Byrne, B. (1998). *The foundation of literacy: The child's acquisition of the alphabetic principle*. Hove, England: Psychology Press.

Carroll, J. B. (Ed.) (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, Massachusetts: The MIT Press.

Cartwright, R. & H. Fraser (forthcoming). *Let your ears do the work! A concept formation approach to teaching pronunciation.*

Celce-Murcia, M., D. Brinton & J. Goodwin (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

Coulmas, F. (1989). What writing means for speech. In F. Coulmas (Ed.), *The writing systems of the world*. Oxford: Basil Blackwells.

Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect: A journal of Australian TESOL*, 21, 44-64.

de Saussure, F. (1986). *Course in general linguistics (transl. Baskin)*. La Salle: Open Court.

Dreyfus, H. (1979). *What computers can't do: The limits of artificial intelligence*. New York: Harper and Row.

Elgin, S. H. (1999). *The language imperative: How learning languages can enrich your life and expand your mind*. Cambridge, Mass: Perseus.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Feynman, R. (1986). *Surely you're joking, Mr Feynman! Adventures of a curious character*. London: Unwin.

Fraser, H. (1992). *The Subject of speech perception: An analysis of the philosophical foundations of the information-processing model of cognition*. London: Macmillan.

Fraser, H. (1997). Phonology without tiers: why the phonetic representation is not derived from the phonological representation. *Language Sciences*, 19, 101-137.

Fraser, H. (2000). *Coordinating improvements in pronunciation teaching for adult learners of english as a second language*. Canberra: Commonwealth of Australia, Department of Education Training and Youth Affairs (Available from http://www-personal.une.edu.au/~hfraser/docs/HF_ANTA_REPORT.pdf).

Fraser, H. (2001). *Teaching pronunciation: A handbook for teachers and trainers*. Sydney: TAFE NSW Access Division.

Fraser, H. (2004a). *Teaching pronunciation: A guide for teachers of English as a second language (CD-ROM, updated)*. Canberra: Commonwealth of Australia, Department of Education Training and Youth Affairs.

Fraser, H. (2004b). Constraining abstractness: Phonological representation in the light of color terms. *Cognitive Linguistics*, 15, 239-288.

Fraser, H. (2006). Helping teachers help students with pronunciation. *Prospect: A journal of Australian TESOL*, 21, 80-94.

Fraser, H. (in press). Categories and concepts in phonology: Theory and practice. In A. Schalley & D. Khlentzos (Eds.), *Proceedings of the International Language and Cognition Conference, sept 2004*. Benjamins.

Fraser, H. (submitted). The sound image as a concept: What adoption of the symbolic thesis means for cognitive phonology. *Cognitive Linguistics*.

Husserl, E. (1960). *Cartesian meditations: An introduction to phenomenology*. The Hague: Martinus Nijhoff.

Ihde, D. (1976). *Listening and voice: A phenomenology of sound*. Athens, Ohio: Ohio University Press.

Langacker, R. (1987). *Foundations of cognitive grammar*. Stanford, California: Stanford University Press.

Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

Lefrancois, G. (1994). *Psychology for teaching (8th ed)*. California: Wadsworth.

Linell, P. (1988). The impact of literacy on the conception of language: The case of linguistics. In R. Saljo, (Ed.), *The written world: Studies in literate thought and action*. Berlin: Springer-Verlag, pp. 41-58.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Macdonald, S. (2002). Pronunciation – views and practices of reluctant teachers. *Prospect: A Journal of Australian TESOL*, 17, 3-15.

McKay, S. L. (2002). *Teaching English as an international language*. Oxford: Oxford University Press.

Merleau-Ponty, M. (1962). The body as expression, and Speech. *Phenomenology of perception*. Routledge Kegan Paul, pp. 174-199.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, Mass: MIT (Bradford).

Nathan, G. S. (1986). Phonemes as mental categories. *Proceedings of the Berkeley Linguistic Society*, 12, 212-223.

Nathan, G. S. (1996). Steps towards a cognitive phonology. In B. Hurch & R. Rhodes (Eds.), *Natural phonology: The state of the art*. Berlin: Mouton, pp. 107-120.

Noble, W. & Davidson, I. (1996). *Human evolution, language and mind: A psychological and archaeological inquiry*. Cambridge: Cambridge University Press.

Ohala, J. J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech*. Cambridge: Cambridge University Press, pp. 258-275.

Olson, D. (1996). Language and literacy: What writing does to language and mind. *Annual Review of Applied Linguistics*, 16, 3-13.

Perkell, J. & Klatt, D. (Eds.) (1986). *Invariance and variability in speech processes*. Hillsdale New Jersey: Lawrence Erlbaum Associates.

Polanyi, M. (1966). *The tacit dimension*. New York: Doubleday.

Premack, D. G. & G. Woodruff (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences*, 1, 515-526.

Rosch, E. (1973). Natural categories. *Cognitive Psychology*, 4, 328-350.

Scarborough, H. S., L. C. Ehri, R. K. Olson & A. E. Fowler (1998). The fate of phonemic awareness beyond the elementary school years. *Scientific Studies of Reading*, 2, 115-142.

Shanon, B. (1993). *The representational and the presentational: An essay on cognition and the study of mind*. New York: Harvester-Wheatsheaf.

Spiegelberg, H. (1982). *The phenomenological movement: A historical introduction*. The Hague: Martinus Nijhoff.

Stillings, N. A., Feinstein, M. H., Garfield, J. L., Rissland, E. L., Rosenbaum, D. A., Weisler, S. E., & Baker-Ward, L. (1987). *Cognitive science: An introduction*. Cambridge, Massachusetts: MIT Press.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Strange, W. (Ed.) (1995). *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press.

Taylor, J. R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.

Taylor, J. R. (2003). *Linguistic categorisation: Prototypes in linguistic theory*. Oxford: Oxford University Press (1st ed. 1989).

Wittgenstein, L. (1958/1974). *Philosophical investigations*. Oxford: Basil Blackwell.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Interlexical Relations in English Stress[1]

FUMIKO KUMASHIRO & TOSHIYUKI KUMASHIRO*
Keio University

ABSTRACT

In this paper, we propose a cognitive, non-reductionist analysis of English stress as it pertains to interlexical relations, based on the usage-based model as proposed by cognitive grammar and on the connectionist interactive activation model. We claim that interlexical relations involved in English stress can felicitously be accounted for by employing actually-occurring expressions as constraints and that precise explication of these relations requires consideration of not only phonological but also semantic factors. In the course of making these claims, we attempt to demonstrate that cognitive grammar, being a usage-based, non-reductionist framework, can accommodate actually-occurring expressions as constraints in a coherent manner and further that the theory can naturally bring semantic factors to bear on phonological analyses, being a non-modular, unificational framework.

KEYWORDS: English stress, cycles, interlexical relations, cognitive grammar, usage-based model, interactive activation model.

* ***Address for correspondence***: Fumiko Kumashiro. Keio University. 4-1-1 Hiyoshi; Kohoku-ku, Yokohama 223-8521; Japan. E-mail: fumiko@kumashiro.org. Toshiyuki Kumashiro. Keio University. 4-1-1 Hiyoshi; Kohoku-ku, Yokohama 223-8521; Japan. E-mail: toshiyuki@kumashiro.org.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I. INTRODUCTION

This paper proposes a cognitive, non-reductionist analysis of English stress as it relates to cycles or INTERLEXICAL RELATIONS. The traditional notion of "phonological cycle" is intended to capture the intuition that words are not merely built by linearly combining morphemes into a string, but can also be derived from independent words, yielding different phonological structures than those which do not contain such independent words. For instance, a monomorphemic word *Mamáronèck* carries primary stress on the antepenultimate syllable, whereas *géneral-ìze*, which is identical to *Mamáronèck* in all relevant respects, is stressed on the initial syllable. This asymmetry can only be ascribed to the fact that *géneral-ìze* is lexically related to *géneral*, which carries stress on the same syllable. More such pairs of examples are provided in (1).[2]

(1)	MONOMORPHEMIC		COMPLEX			
	Saskátchewàn	vs.	**oxýgen-àte*	vs.	*óxygen-àte*	(cf. *óxygen*)
	ínventòry	vs.	**ínfirmàry*	vs.	*infírmary*	(cf. *infírm*)
	mércantìle	vs.	**pércent-ìle*	vs.	*percént-ìle*	(cf. *percént*)

Note that not only the primary stress but also the secondary one can be affected by the stress patterns of lexically-related words. Observe the examples in (2). A monomorphemic word *àbracadábra* in (2)a, for instance, carries secondary stress on the initial syllable. In contrast, a complex word *orìginál-ity* in (2)b carries it on the second syllable despite it having a comparable syllable structure. Notice that the complex word in question has a related independent word that is stressed on the same syllable, i.e. *oríginal*; only this fact could possibly explain the contrast observed here.[3]

(2)a.	MONOMORPHEMIC	b. COMPLEX			
	àbracadábra	**òriginál-ity*	vs.	*orìginál-ity*	(cf. *oríginal*)
	pàraphernália	**ìconoclást-ic*	vs.	*icònoclást-ic*	(cf. *icónoclàst*)
	Kìlimanjáro	**ànticipát-ion*	vs.	*antìcipát-ion*	(cf. *antícipàte*)

Therefore, one may be inclined to conclude that those words which are derived from other related independent words are stressed exactly where the independent words are stressed. The facts,

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

however, are more complicated, and the above conclusion does not always obtain. That is, there are not a few words which carry stress on the syllables other than the ones on which their related independent words are stressed. Those words typically behave, at least with respect to stress, as if they were monomorphemic words. Observe the examples in (3) below:

(3)a.	*ínfer-ence*	vs.	**infér-ence*	(cf. *infér*)
	réside-ence	vs.	**resíde-ence*	(cf. *resíde*)
	prótest-ant	vs.	**protést-ant*	(cf. *protést*)
b.	*parént-al*	vs.	**párent-al*	(cf. *párent*)
	solíd-ify	vs.	**sólid-ify*	(cf. *sólid*)
	orígin-àte	vs.	**órigin-àte*	(cf. *órigin*)
	demócrat-ìze	vs.	**démocrat-ìze*	(cf. *démocràt*)
c.	*ím-pious*	vs.	*im-píous*	(cf. *píous*)
	èlemént-ary	vs.	**élement-àry*	(cf. *élement*)

Examples in (3)a carry primary stress on the initial syllable, although their related independent words are stressed on the second. The converse is true for examples in (3)b: derived words are primarily stressed on the second syllable, whereas the source words are stressed on the initial. Some additional cases of similar mismatch are found in (3)c.

Any phonological theory, therefore, faces the extremely difficult task of providing a mechanism capable of encoding the interlexical relations as illustrated by (1) and (2) above and that of explaining at the same time why the relations do not hold for the words in (3). In the current paper, we attempt to propose an analysis that can in principle be successful with the tasks defined above, based on the usage-based model (Langacker, 1988) within the framework of cognitive grammar (Langacker, 1987, 1990, 1991, 1999), and on the connectionist interactive activation model (Elman & McClelland, 1984; Rumelhart & Zipser, 1985; Waltz & Pollack, 1985). More specifically, the current paper aims to make the following two claims: (i) interlexical relations involved in English stress can felicitously be accounted for by employing an actually-occurring expression as a constraint; and (ii) precise explication of interlexical relations requires consideration of not only phonological but also semantic factors. In the course of making these

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

claims, we also intend to demonstrate the following with respect to the framework: (i) cognitive grammar, being a usage-based, non-reductionist framework, can accommodate actually-occurring expressions in a coherent manner without employing ad-hoc mechanism; (ii) the theory can naturally accommodate semantic factors in phonological analyses, being a non-modular, unificational framework that employs in phonological analyses only those theoretical constructs which have already been proposed elsewhere for semantic analyses; and (iii) the theory, whose focus has primarily been on semantic analyses, is capable of offering a framework in which phonological phenomena can be successfully accounted for (cf. Farrell, 1990; Kumashiro, F., 2000; Kumashiro, T., 1990; Rubba, 1993).

The organization of the current paper is as follows. In Section II, the model, principles, and representations that would form the basis of the proposed analysis will be presented. Section III illustrates how cognitive grammar can handle prototypical cases of interlexical relations. Section IV deals with exceptional cases. Section V offers comparison with a comparable analysis in optimality theory.

II. MODEL, PRINCIPLES, AND REPRESENTATIONS

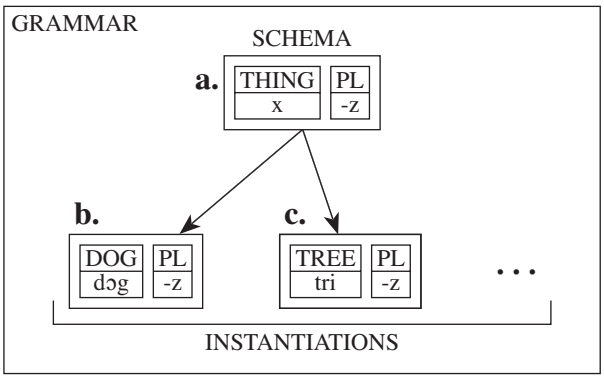
II.1. The Usage-Based Model

The model to be used to account for the data related to English stress in question is based on the usage-based model, proposed within the general framework of cognitive grammar. The theory views grammar as a structured inventory of conventional linguistic units. That is, it is essentially a "bottom-up, non-reductionist, maximalist" approach, in which the grammar is viewed as storing every conventionalized expression (INSTANTIATION) as well as generalizations (SCHEMAS) that may have been schematized by language users from actually-occurring concrete expressions.[4] Therefore, there is no fundamental difference in theoretical status between actually-occurring expressions and generalizations, which only differ in terms of degree of specificity. This situation is illustrated in Fig. 1 (adapted from Langacker, 1988: 131). Provided in Fig. 1b is the representation for a plural noun *dogs* and in Fig. 1c that for *trees*. Notice that the words are

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

morphologically complex and can be decomposed into parts: *dogs* into *dog* and *-s*, and *trees* into *tree* and *-s* (in the diagram, these expressions are placed in separate boxes). Furthermore, every linguistic expression is BIPOLAR and can be separated into the SEMANTIC and the PHONOLOGICAL POLE: *dog* into the semantic pole [DOG] and the phonological pole [dɔg], *tree* into [TREE] and [tri], and the plural suffix into [PL] and [-z] (capital letters are used to represent the semantic pole).[5] As there are numerous other plural nouns in the grammar (symbolized by "elliptical" three dots in the diagram), which share the same internal structure, it is reasonable to assume that the language user has extracted a SCHEMA (given in Fig. 1a) which has exactly the same internal composition as *dogs* and *trees*, but has a noun stem whose semantic and phonological poles are characterized only SCHEMATICALLY.[6]

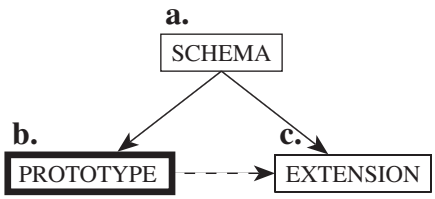
Fig. 1: Organization of Grammar



Furthermore, conventionalized linguistic units are not merely "listed" in the grammar but they are also "structured". A typical structure takes the form described in Fig. 2 (adapted from Langacker, 1988: 140). For any type of linguistic category, there is a PROTOTYPE, which is a central member of the category.[7] There are other peripheral members of the category, which resemble or overlap with the prototype to various degrees (EXTENSIONS). One can usually posit a SCHEMA which has all and only the properties common to both the prototype and the extensions and thus is schematic with respect to them.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Fig. 2: Categorization

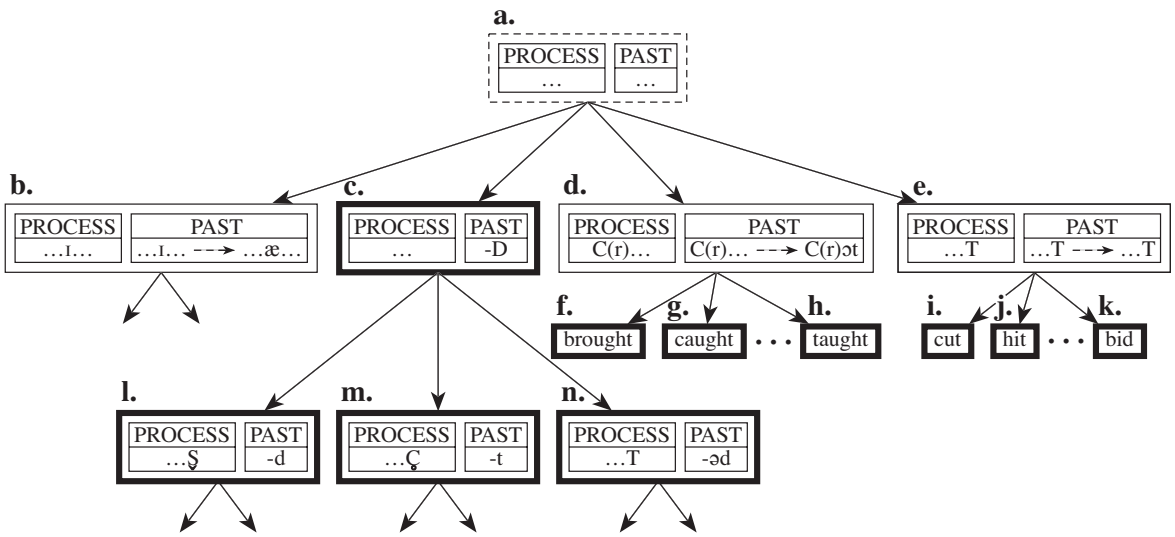


Furthermore, the grammar, in addition to containing those entries or nodes, specifies the relations holding among them. There are two types of such relationships: SCHEMATIZATION and EXTENSION. The relation of extension (symbolized by a dashed-line arrow) implies some conflict in specifications between the basic and extended values; hence [A] - - → [B] indicates that [B] is incompatible with [A] in some respect, but is nevertheless categorized by [A]. The relation of schematization, on the other hand, holds between a schema and a structure that elaborates or instantiates the schema. Symbolized by a solid-line arrow, e.g. [A] ⟶ [B], the relationship amounts to one of specification: [B] conforms to the specification of [A], but is characterized with finer precision and detail. In this fashion, conventionalized units form massively connected networks for each relevant cognitive domain.

Let us now examine what a typical network looks like. Described in Fig. 3 is a network for the English past-tense morpheme (adopted from Langacker, 1988: 155). Described in Figs. 3l-n at the bottom are the most frequently used past-tense forms, representing the regular patterns: [-d], [-t], and [-əd] suffixations. Notice these schemas contain not only the structures for the suffixes themselves, but also the schematic characterizations of the stem verbs. For example, the schema for the [-d] suffix (Fig. 3l) contains a schematically characterized verb stem, of which the phonological pole only stipulates that the stem ends in [ʂ] (a voiced segment), while the semantic pole only has the specification "PROCESS", which is the highly schematic semantic value common to all verbs, encompassing both stative and perfective verbs. Likewise, the schema for past-tense forms with the suffix [t] stipulates that the stem ends in [ʂ̥] (a voiceless consonant) at the phonological pole; and the schema for [-əd], that the stem ends in [T] (an alveolar stop neutral with respect to voicing).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Fig. 3: English Past Tense



Described in Fig. 3c is the higher schema extracted from the above-mentioned three concrete schemas. Being schematic over a wide range of verb stems, the phonological pole of the stem is almost vacuous and simply stipulates that there be some phonological content ([…]). The phonological pole of the suffix itself is simply specified as [-D] (phonological content neutral with respect to voicing and the presence of a schwa). Depicted in Fig. 3b is the schema encompassing verbs whose past-tense forms involve the substitution of some vowel by [æ] (e.g. *rang*, *sank*, *swam*, *sat*).[8] Described in Fig. 3d is the schema for past-tense forms containing [ɔt] at the end, and in Fig. 3e is the schema for past-tense forms that end in an alveolar and are identical to the infinitival forms. Recall that cognitive grammar is a "maximalist" approach, which lists not only schemas but also expressions that actually occur in the grammar; this is why *brought*, *caught*, *taught*, *cut*, *hit*, *bid*, and abbreviated many others are listed under their respective schemas. Finally, in Fig. 3a is the super schema covering all the past-tense forms. Its wide-range applicability makes it highly schematic; the phonological and the semantic pole simply stipulate that there be some content ([…]).

Also described in Fig. 3 is the cognitive salience of each schema and each actually-occurring expression. The schemas in Figs. 3l, m, and n for the prototypical plural forms with the suffixes

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

[-d], [-t], and [-əd], respectively, and the higher-level schema in Fig. 3c for these lower schemas are accorded high cognitive salience (symbolized by bold-line boxes) because they represent most common types of past-tense forms. The actually-occurring expressions (such as *brought* in Fig. 3f and *cut* in Fig. 3i) are given high cognitive salience as well because they are concrete expressions and occur more frequently than regular verbs. The super schema in Fig. 3a enjoys only low salience (symbolized by a dashed-line box) because it is almost void of semantic and phonological contents. All the other schemas are accorded an intermediate degree of salience.

It should be apparent from the diagram that cognitive grammar does not rely on such theoretical constructs as rules and rule ordering, and that all the generalizations are instead stated in the form of schemas. This is required by the highly restricted principle of the theory, namely the CONTENT REQUIREMENT, which permits in grammar only "(1) phonological, semantic, or symbolic structures that actually occur in linguistic expressions; (2) schemas for such structures; and (3) categorizing relationships involving the elements in (1) and (2)" (Langacker, 1987: 53-54). Therefore, the well-formedness of an expression cannot be determined by whether it can be generated by rules; it is instead determined by whether it is categorized by a schema. Relevant rinciples for determining such well-formedness are discussed below.

II.2. Well-formedness principles

Langacker (1988: 153) proposes the principles in (4) as a working hypothesis for determining the well-formedness of an expression:

(4) a. Uniqueness

When an expression is assessed relative to a grammatical construction, a single node (from the network representing the construction) is activated for its categorization; if this "active node" is schematic for the expression, the latter is judged well-formed (conventional).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

b. Selection

The likelihood that a given node will be chosen as the active node for categorizing a target expression correlates positively with its degree of entrenchment and cognitive salience, and negatively with its "distance" from the target, i.e. how far the target diverges from it by elaboration or extension.

Note that the selection principle in (4)b entails that the notion of well-formedness is not categorical but gradient. Furthermore, the uniqueness and selection principles in (4) are compatible with the connectionist interactive-activation model of language processing. However, in order to make the two principles more harmonious with the model, we propose the following revisions, based on proposals made by T. Kumashiro (1990) and F. Kumashiro (2000):

(5) a. Access

When a given candidate expression is assessed relative to a certain subpart of the grammar, i.e. a function, units (from the network representing the subpart) that categorize the expression are activated and sanction the expression.

b. Activation

The total "activation", i.e. conventional motivation/sanction, of a candidate expression is the sum of the activation values obtained from all of the categorizing units. Each such value correlates positively with the degree of entrenchment and cognitive salience of the unit, and negatively with the expression's "distance" from the unit, i.e. how far it diverges from its categorizing unit by elaboration or extension.

c. Uniqueness

When there are multiple candidate expressions, all but the one with the highest activation value are deactivated.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

d. Well-Formedness

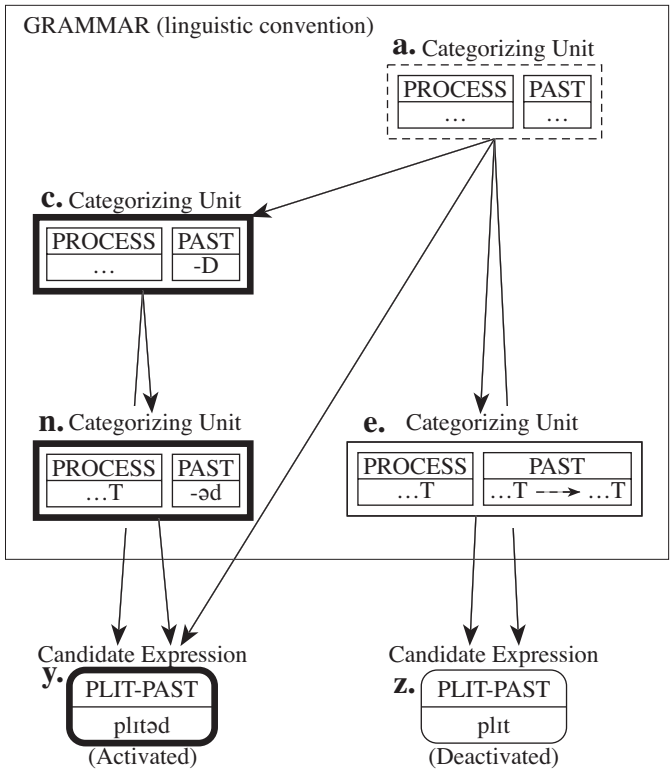
The degree of well-formedness of a candidate expression correlates with its final activation value.

Note that Langacker's proposal in (4) and ours in (5) are essentially equivalent, despite ostensible differences. In Langacker's formulation, the task of the speaker is to choose among "schemas", whereas in ours, the selection is made among possible "candidates".[9]

Let us provide a specific example to illustrate these principles. Part of the network in Fig. 3 that represents English past-tense formation is provided in Fig. 4.[10] Given a novel verb such as *plit* [plɪt], several candidate expressions with the function of its past-tense form are conceived: *plitted* [plɪtəd], *plat* [plæt], *plaught* [plɔt], and *plit* [plɪt]. In what follows, the competition between the two most plausible candidates, i.e. *plitted* [plɪtəd] and *plit* [plɪt], is examined. The candidate expression *plitted* in Fig. 4y is categorized as an instantiation not only of the lower-level categorizing unit in Fig. 4n but also of the higher-level ones in Figs. 4c and a. Then, in accordance with the access principle in (5)a, all the three categorizing units are activated and sanction *plitted*, but to different degrees. According to the activation principle in (5)b, the total activation value of *plitted* is the sum of the activation values obtained from the categorizing units in Figs. 4n, c, and a. Likewise, the candidate expression *plit* in Fig. 4z is categorized as an instantiation not only of the lower-level schema in Fig. 4e but also of the higher-level one in Fig. 4a. Both units are activated and sanction *plit*, and the total activation value of *plit* is the sum of those obtained from the categorizing units in Figs. 4e and a.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Fig. 4: Past-Tense Form of Novel Verb *plit*



Next, we need to compare the total activation value of *plitted* in Fig. 4y against that of *plit* in Fig. 4z, in order to determine which candidate expression is the most activated. Both candidate expressions are categorized by the categorizing unit in Fig. 4a; from the unit, they obtain effectively the same activation values because the difference in the distance from the unit is negligible, if any. As for other categorizing units, the distance from the unit in Fig. 4n to the candidate expression in Fig. 4y the unit categorizes as an instantiation is essentially equal to that between the unit in Fig. 4e and the candidate expression in Fig. 4z the unit categorizes because the phonological poles of the stems of the two units are identical ([…T]). Therefore, the distance criterion of the activation principle in (5)b does not play a role in determining which is higher of the value the candidate expression in Fig. 4y obtains from the categorizing unit in Fig. 4n and that the candidate in Fig. 4z obtains from the unit in Fig. 4e. Instead, the decision hinges on the cognitive salience criterion of the principle. As can be observed in the networks in Figs. 3 and 4, the

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

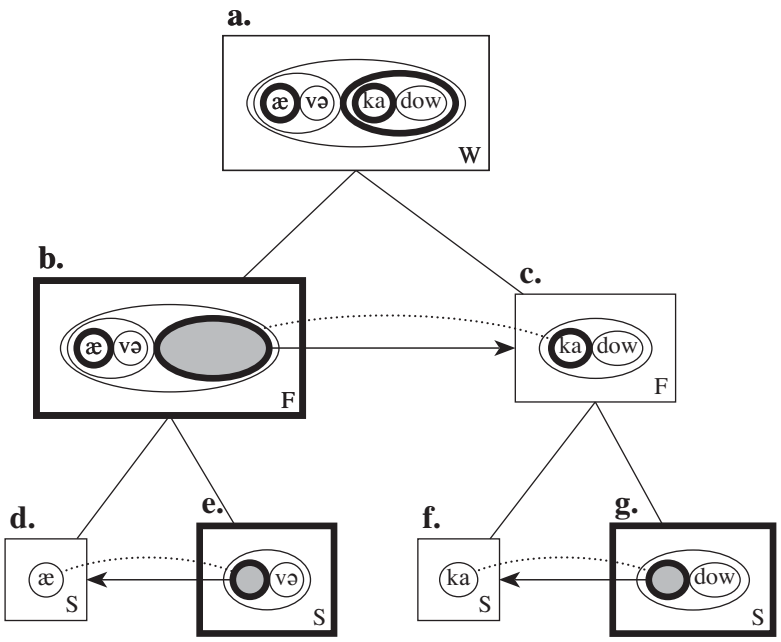
categorizing unit in Fig. 4n, as a prototypical schema representing a regular pattern, is far more salient than that in Fig. 4e. Therefore, the candidate expression in Fig. 4y obtains a higher activation value from the categorizing unit than the value the candidate in Fig. 4z obtains from the unit in Fig. 4e. Moreover, the candidate expression in Fig. 4y obtains an additional activation value from the categorizing unit in Fig. 4c. Therefore, the candidate has a higher total activation value than the one in Fig. 4z. As a result, the uniqueness principle in (5)c allows the candidate to remain active, but deactivates the one in Fig. 4z. In sum, the well-formedness principles in (5) successfully predict that as the past-tense form of a novel verb *plit*, *plitted* is judged well-formed, but *plit* is not.

II.3. Prosodic representations

Before presenting an analysis of English stress involving complex words, it is appropriate to present the cognitive grammar representation of prosodic structure at the word level. Described in Fig. 5 are different levels of prosodic representation for the noun *àvocádò*. Depicted in Figs. 5d-g at the bottom is the syllable-level organization, where *àvocádò* is shown to be comprised of four syllables: [æ], [və], [ka], and [dow].[11] In this example, the initial syllable [æ] is strong with respect to the antepenult [və]. A strong syllable is considered AUTONOMOUS as it can occur "in full, unreduced form—approximately as if it were pronounced in isolation" (Langacker, 1987: 331). A weak syllable, on the other hand, is DEPENDENT because it is "compressed along such phonetic parameters as time, amplitude, and pitch range" (loc. cit.), and to be implemented with such properties, it "must be pronounced in combination with an autonomous one" (loc. cit.). That is to say, the prosodic representation for a weak syllable at the lexical level should include a schematic reference to a strong one as part of its inherent characterization. The diagram for the weak syllable [və] in Fig. 5e thus includes the phonological specifications for two syllables: the elaborated, weak syllable on the right (a circle is used to represent syllablehood) and the schematic, strong syllable on the left (a bold line is used to represent prosodic prominence). In a similar fashion, the phonological characterization of the weak syllable [dow] includes the specifications for a schematic strong syllable and an elaborated weak syllable.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Fig 5: Prosodic Structure of *àvocádo*



At a higher level of organization illustrated in Fig. 5b, the two syllables [æ] and [və] are integrated to form a foot. This integration is effected by the equation (symbolized by a dotted line) of overlapping phonological specifications at the syllable level—i.e. the schematic syllable in Fig. 5e is equated with the elaborated one in Fig. 5d. As a result, the schematic syllable gets elaborated by the corresponding, elaborated syllable. In the same manner, the strong penult [ka] and the weak ultima [dow] are integrated into the foot [ká.dòw]. Notice that there is discrepancy in phonetic prominence between the two feet, i.e. [ǽ.və] and [ká.dòw], comparable to that between the strong syllables [æ] and [ka] on one hand and the weak syllables [və] and [dow] on the other: [ká.dòw] is strong vis-à-vis [ǽ.və]. As in the case of syllables, the phonological specification for a weak foot includes the schematic characterization of a strong foot. The two feet are integrated to form a word at the next higher level, in the same manner the foot-level integration is effected, as illustrated in Fig. 5a.

As the result of these integrations at the foot and the word level, the prosodic representation of the word [æ̀.və.ká.dòw] as a whole involves three layers: the syllable, foot, and word levels

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

(from inside out). Furthermore, each layer specifies the relative prominence of their components. Notice that specifying the level of each component and its relative prominence at the given level is sufficient to determine the relative prominence of all syllables at the word level: the most prominent syllable at the word level is the syllable that is strong at the foot level and is contained in the strong foot, and the least prominent syllable is the one that is the weak syllable of the weak foot. Thus, the penult [ka] is the most prominent one at the word level because it is strong within the foot [ká.dòw], which is in turn strong within the word [æ.və.ká.dòw]. The antepenult [və] is the least prominent syllable, as it is weak within the foot [ǽ.və], which is weak within the entire word. The initial syllable [æ] is of intermediate prominence, for it is contained in the weak foot [ǽ.və] although it is strong itself within the foot. The final syllable [dow] is also accorded intermediate prominence because it is a weak syllable itself, albeit contained within the strong foot [ká.dòw].

III. INTERLEXICAL RELATIONS IN COGNITIVE GRAMMAR

Now that we have discussed the usage-based model, the well-formedness principles, and the prosodic representations, we are ready to present the analysis of the stress patterns of lexically complex words in English. Described in Fig. 6 is the network of schemas relevant for the determination of the phonological well-formedness of the lexically complex verb *géneral-ìze*. The representation for the verb itself is provided in Fig. 6c. Notice that the representation is comprised of two parts, the semantic pole (at the top) and the phonological pole (at the bottom), which stand in a symbolic relationship (represented by a dotted line).[12] Provided at the phonological pole are the phonological specifications: the word consists of the strong foot [ǰɛ.nə.rə] and the weak foot [layz]. The former consists of the strong syllable [ǰɛ] and the weak syllables [nə] and [rə], while the latter contains only one syllable, [layz]. Just as the phonological pole has complex internal structure, so is the semantic structure: the semantic specifications for *géneralìze* include those for the root word *géneral* and the affix *-ize*; and the existence of specifications for the root word *géneral* in the complex word *géneralìze* is readily recognized, i.e. the language user is very likely to be aware of the existence of such substructure. Sketched in Fig. 6a is the adjectival lexical unit *géneral*. This

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

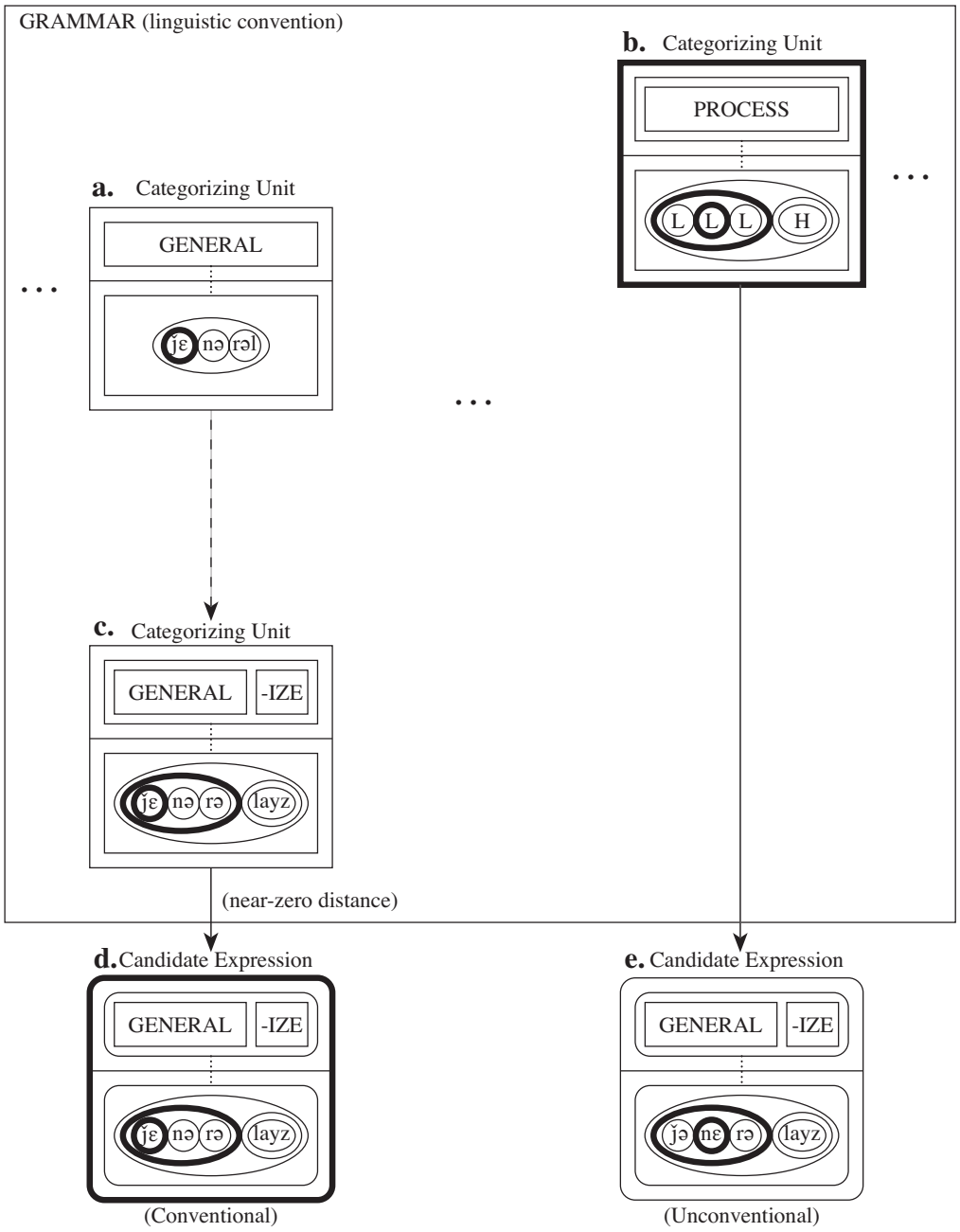
unit categorizes *géneralìze* in Fig. 6c as an extension from it with some negligible conflict in specification, apart from the conflict caused by the addition of the semantic and phonological structure for the suffix *-ize* (thus a dashed-line arrow, not a solid-line one for instantiation, is used to connect the two nodes in the diagram).[13] Sketched in Fig. 6b is the schema extracted from verbs which have an initial strong foot consisting of three light syllables with the middle one strong, combined with a final weak foot comprised only of one heavy syllable.[14]

All of the three nodes explained so far—for the lexical units *géneralìze* in Fig. 6c and *géneral* in Fig. 6a, and for the schema in Fig. 6b—are all included in the grammar as "conventionalized" linguistic expressions. What is to be noted here is that the lexical units and the schema are on a equal footing and both serve as categorizing units, which is how interlexical relations are coded in the current analysis. However, they differ in terms of prominence. The schema in Fig. 6b is to be considered more cognitively salient than either of the nodes for the lexical units *géneralìze* in Fig. 6c and *géneral* in Fig. 6a for its ease of activation (the cognitive salience of the schema is symbolized by the use of a bold line for the enclosing box).

Fig. 6 also includes two candidate expressions at the bottom, i.e. the conventional candidate *géneralìze* in Fig. 6d and the unconventional candidate in Fig. 6e. The former is categorized by the node for *géneralìze* in Fig. 6c as an instance, and the latter by the schema in Fig. 6b. When a speech-act participant attempts to choose one of the two candidates as the prevailing candidate, s/he assesses them against the grammar, and does so against the above three nodes in particular, according to the well-formedness principles in (5). However, the process of selecting the prevailing candidate is "trivial" in the case of a conventionalized expression such as *géneralìze*, because the grammar already contains it as a conventional expression (Fig. 6c) and the cognitive distance between the candidate expression in Fig. 6d and the conventionalized expression is completely negligible, if not zero. As a result, the candidate expression conforming to the conventional pattern (Fig. 6d) always wins out over ones that do not (e.g. Fig. 6e) without exceptions.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Fig. 6: *géneralìze*



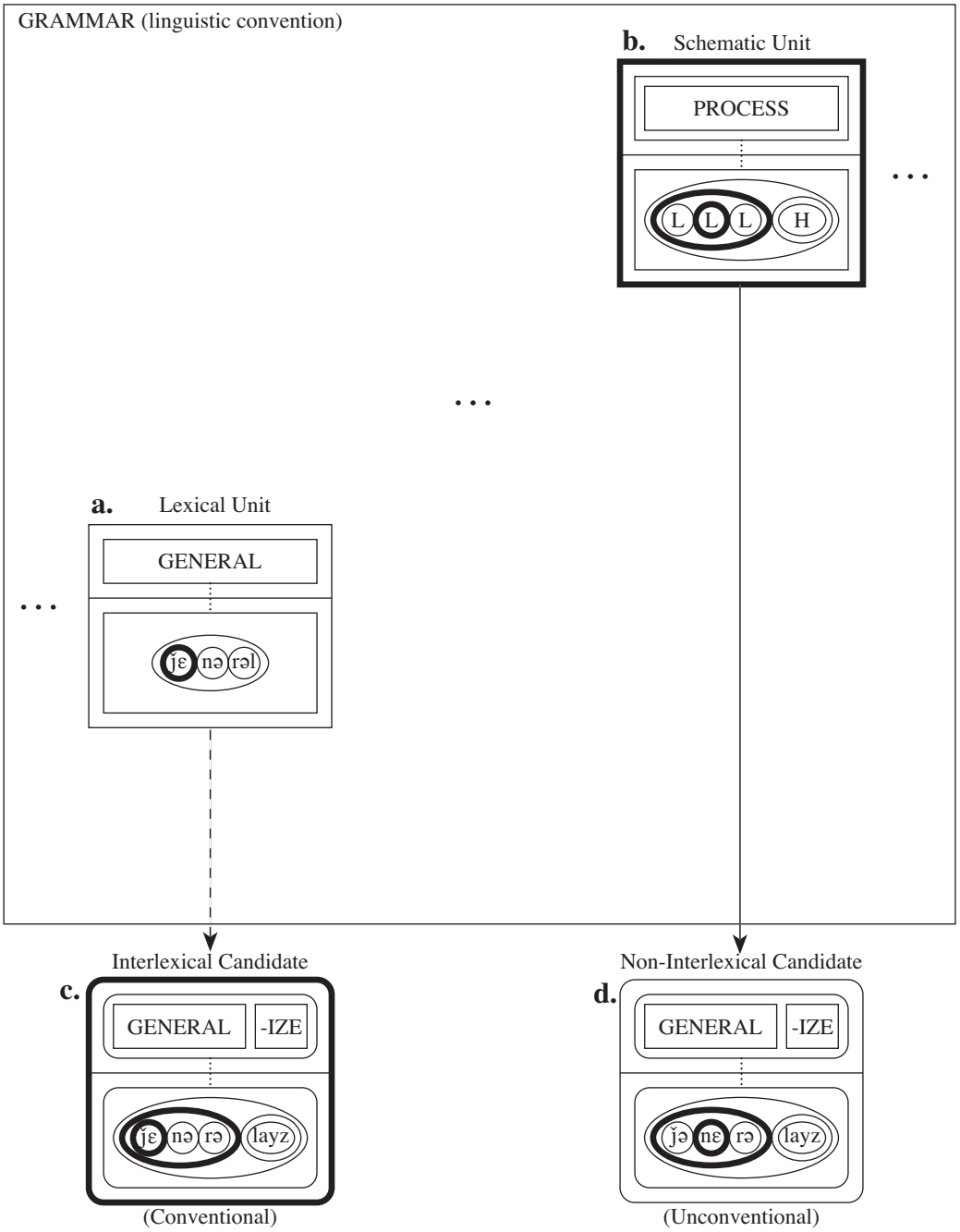
Therefore, if one wishes to examine the "predictability", in the commonly-used sense of the word, of the network respect to the well-formedness of a given expression, one must assume that the expression in question is not conventionalized. One can easily create this situation by simply

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

removing the node for the expression from the grammar. The case of *géneralìze* with this modification is described in Fig. 7. Here the process of selecting the prevailing candidate is no trivial and requires dynamic calculations of phonological distances between the candidates and their categorizing units. First, the node for the adjective *géneral* in Fig. 7a categorizes the conventional INTERLEXICAL CANDIDATE *géneralìze* in Fig. 7c as an extension with some negligible conflict other than what is caused by the suffixation with *-ize*. The negligible conflict in question is in the phonological pole: the ultima of [ǰɛ.nə.rəl], i.e. [rəl], is a heavy syllable having [l] as the coda, whereas the corresponding syllable in [ǰɛ.nə.rə.layz], i.e. the penult [rə], lacks the coda. However, the degree of extension caused by the conflict is minimal: the structure of a tri-syllabic foot constituting the phonological pole of *géneral* is found intact in *géneralìze*. Next, the schema in Fig. 7b categorizes the unconventional NON-INTERLEXICAL CANDIDATE *genéralìze* in Fig. 7d. This stress schema is extracted from those verbs which have an initial strong foot consisting of three light syllables with the middle one strong and a final weak foot comprised only of one heavy syllable. It is reasonable to posit a schema like this in light of a number of verbs which conform to the above phonological specifications (e.g. *inóculàte, accómmodàte, affíliàte*). In order to determine which candidate should prevail, we need to compare the activation value the interlexical candidate *géneralìze* in Fig. 7c obtains from the lexical unit *géneral* in Fig. 7a against the value the non-interlexical candidate *genéralìze* in Fig. 7d receives from the schematic unit in Fig. 7b, according to the activation principle in (5)b. In terms of cognitive distance, the distance between the former pair of nodes is determined to be far smaller than that between the latter pair, because the specifications in the lexical unit *géneral* in Fig. 7a for the internal structures of the semantic and the phonological pole are more elaborate than those in the schematic unit in Fig. 7b. In terms of salience, the schematic unit in Fig. 7b is considered more salient than the lexical unit in Fig. 7a because it directly categorizes a large number of expressions. However, the lexical unit's far smaller distance from the interlexical candidate in Fig. 7c more than compensates its lesser degree of salience; as a result, the interlexical candidate obtains a higher activation value from the lexical unit in Fig. 7a than what the non-interlexical candidate in Fig. 7d does from the schematic unit in Fig. 7b.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Fig. 7: *géneralize* (Hypothetical)



Therefore, the network would correctly predict that when *géneralize* and *genéralize* are put in competition, as alternative pronunciations for an unconventional word would be in the course of

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

time, the former should win out. In sum, the network predicts that, provided a high degree of both phonological and semantic decompositionality, the stress pattern analogous to the root word, not one analogous to the monomorphemic word, should prevail.

IV. NON-INTERLEXICAL CASES

We observed that a high degree of phonological and semantic decompositionality results in the prevalence of the interlexical candidate exhibiting a stress pattern comparable to the root word. Notice, however, that this statement entails that a lesser degree of either phonological or semantic decompositionality would affect the distances between the categorizing units and the candidate expressions, thereby changing the dynamics of the network and the predictions it makes.

IV.1. Phonological Opacity

To see a case of decreased phonological decompositionality, or phonological opacity, affecting the stress pattern, let us consider the case of *solídify*, illustrated in Fig. 8. In this case, the non-interlexical *solídify* prevails over the interlexical *sólidify* because of the phonological opacity of the latter. There are two candidates in Fig. 8: the interlexical yet unconventional candidate *sólidify* in Fig. 8c and the non-interlexical but conventional candidate *solídify* in Fig. 8d. The latter candidate is categorized by the schematic unit in Fig. 8b, which is identical to the one in Fig. 7b. The former is categorized by the lexical unit *sólid* in Fig. 8a as an extension with some noticeable conflict. Notice here that the distance between the lexical unit *sólid* in Fig. 8a and the unconventional interlexical candidate *sólidify* in Fig. 8c is considered greater than that between the schematic unit in Fig. 8b and the conventional non-interlexical candidate *solídify* in Fig. 8d, because the phonological decomposability of the interlexical candidate *sólidify* in Fig. 8c into *sólid* in Fig. 8a is lower than that of *géneralìze* in Fig. 7c into *géneral* in Fig. 7a. The phonological opacity observed here stems from conflict in foot-internal structure. In the *géneral*/*géneralìze* case, the adjective *géneral* (Fig. 7a) forms a single foot consisting of three syllables ([ʃɛ.nə.rəl]), and in

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

the verb *géneralìze* (Fig. 7c), a comparable tri-syllabic foot ([ǰɛ.nə.rə]), is found. In the *sólid*/*sólidifŷ* case, on the other hand, the adjective *sólid* (Fig. 8a) forms a bi-syllabic foot ([sa.ləd]), but in *sólidifŷ* (Fig. 8c), the comparable foot contains three syllables ([sa.lə.də]). This greater distance between the lexical unit *sólid* in Fig. 8a and the interlexical unit *sólidifŷ* in Fig. 8c results in the prevalence of the non-interlexical candidate *solídifŷ* in Fig. 8d, which follows the stress pattern of comparable monomorphemic words.

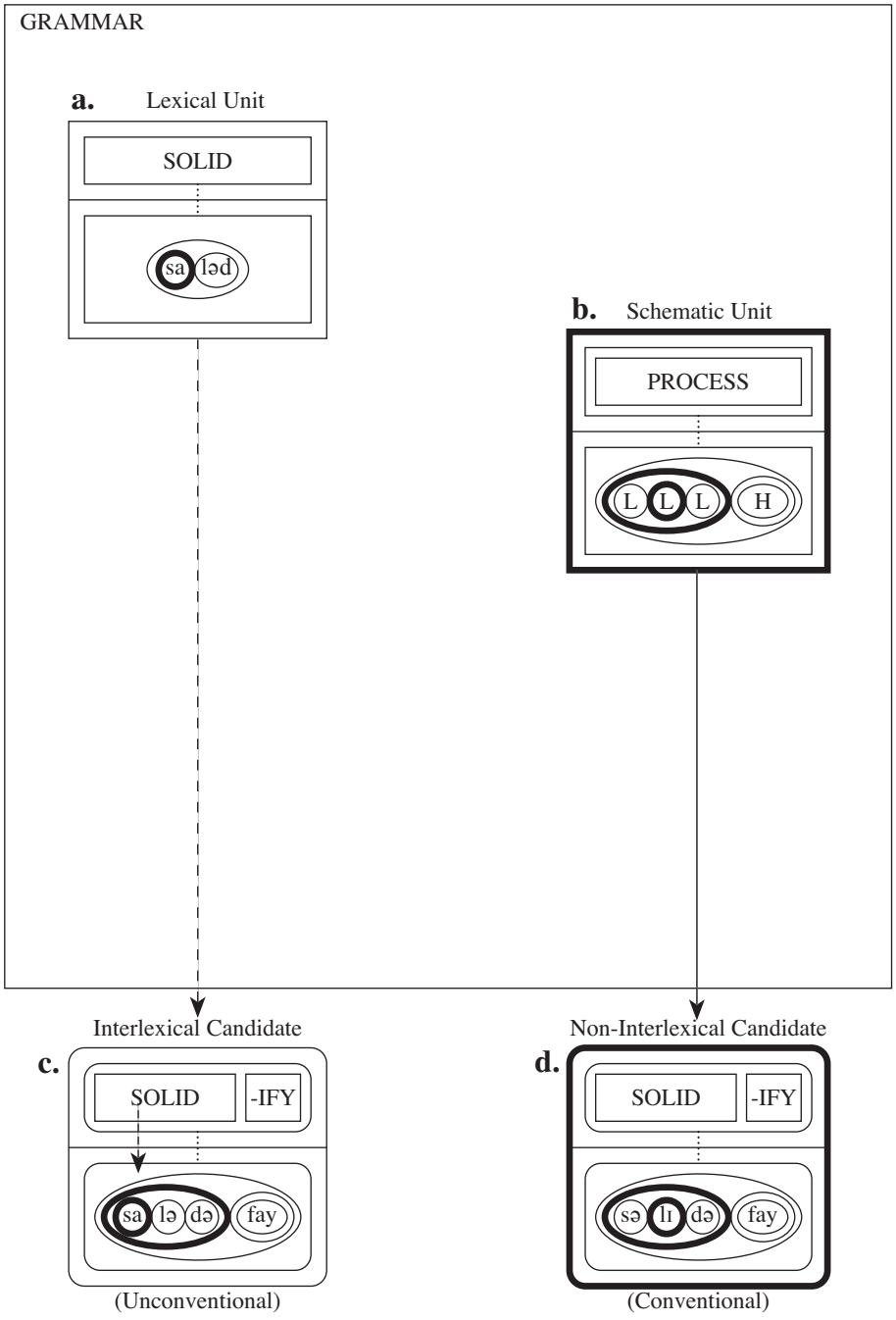
In sum, when phonological opacity is observed, the non-interlexical candidate whose stress pattern is distinct from the root word but analogous to comparable monomorphemic words is predicted to be prevalent, in contrast to the *géneralìze* case illustrated in Fig. 7, which exhibits the opposite pattern, involving a higher level of phonological decomposability. More examples involving phonological opacity are provided in (6). Note that in all these examples, the root words form bi-syllabic feet, but the corresponding feet in the unconventional interlexical candidates are all tri-syllabic.

(6)

		TRISYLLABIC	BISYLLABIC
ínfer-ence	vs.	**infér-ence*	(cf. *infér*)
réside-ence	vs.	**resíde-ence*	(cf. *resíde*)
prótest-ant	vs.	**protést-ant*	(cf. *protést*)
parént-al	vs.	**párent-al*	(cf. *párent*)

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Fig. 8: Phonological Opacity (*sólidify*)



© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

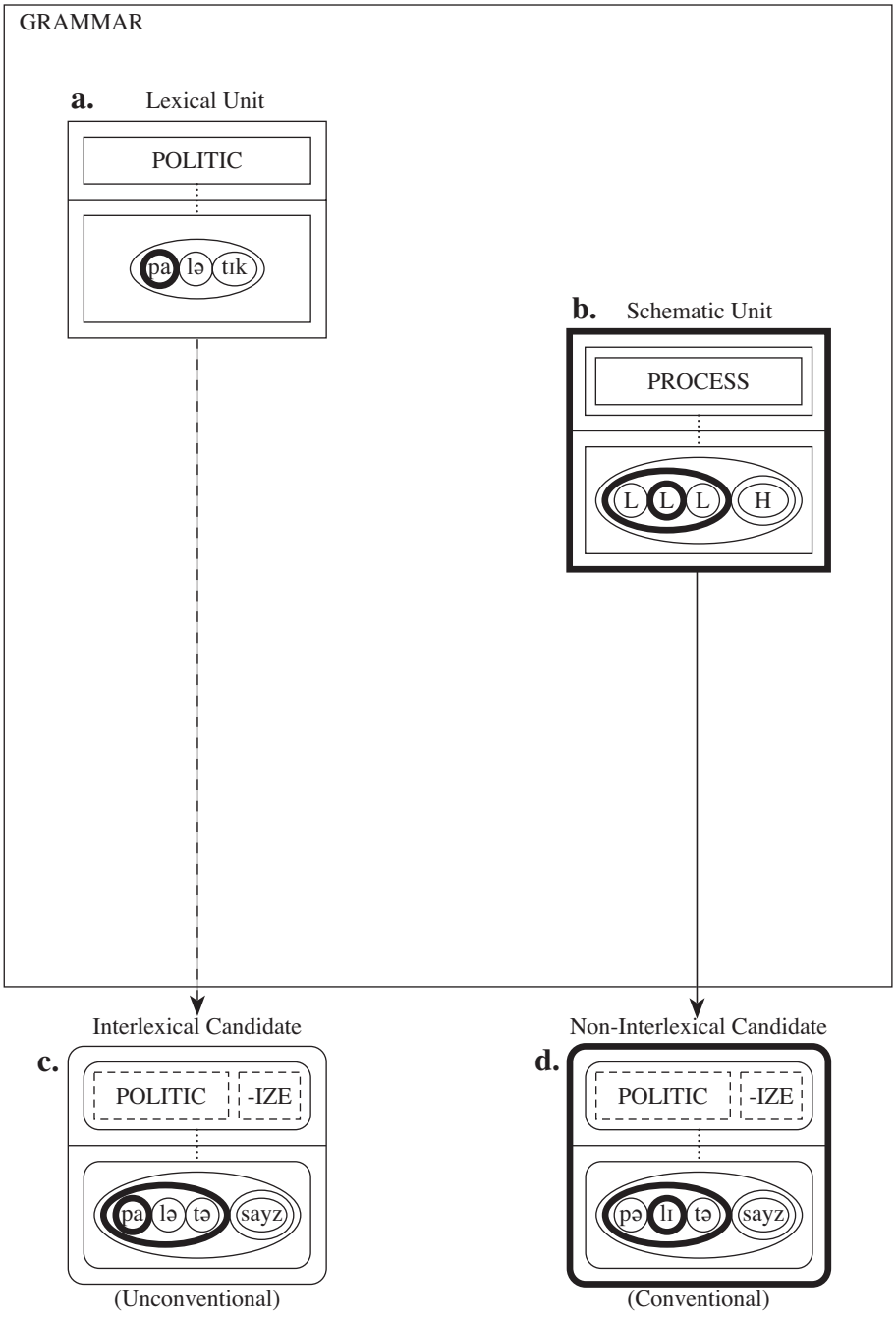
IV.2. Semantic opacity

Next, let us examine the case of decreased "semantic" decompositionality, or semantic opacity, resulting in the prevalence of the phonological structure analogous to comparable monomorphemic words. Consider the case of *polítícize*, depicted in Fig. 9. In this case, the non-interlexical *polítícize* prevails over the interlexical *póliticize* because of semantic opacity. The configuration of the network for *polítícize* in Fig. 9 is exactly the same as that for *solídify* in Fig. 8. However, there is some difference in the nature of the cause for the greater distance between the lexical unit (*pólitic* in Fig. 9a) and the interlexical candidate (*póliticize* in Fig. 9c); the cause is the decreased "semantic", rather than "phonological", decompositionality. That is, the semantic pole of *pólitic*, which is used to describe characteristics not necessarily related to politics, is not as readily recognizable in that of *polítícize*, as *géneral* is in *géneralize*. More similar examples involving decreased semantic decomposability are provided in (7):

(7) *eleménT-ary* vs. **élement-ary* (cf. *élement*)
orígin-ate vs. **órigin-ate* (cf. *órigin*)

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Fig. 9: Semantic Opacity (*políticìze*)



© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

V. INTERLEXICAL RELATIONS IN OPTIMALITY THEORY

Since Chomsky & Halle (1968), many generative phonologists attempted to explain interlexical relations between independent words, as observed in (1) above, by deriving one from another using the notion of "cycle".[15] In this type of "cyclical" approaches, the effect of the metrical structure of the smaller word on that of the larger is automatic, for the latter structure is literally "built from" the former. In terms of the representation systems employed, there have been basically two different approaches; namely, the tree theory (Hayes, 1982, 1984; Kiparsky, 1979, 1982; Liberman & Prince, 1977) and the grid theory (Halle & Vergnaud, 1987; Prince, 1983; Selkirk, 1984). The most recent addition to this tradition is optimality theory, which abolishes rules and derivation, and instead relies on simultaneous evaluation of competing constraints. In what follows, we will see how interlexical relations are handled in the theory.

In optimality theory, interlexical relations are formulated as constraints on the correspondence between one output and another (Benua, 1997, 2000, 2004). Such an analysis is illustrated in (8) (Benua, 1997: 27):

(8) Transderivational (Output-Output) Correspondence

	OO-Correspondence	
[root$_i$]	→	[root$_i$ + affix]
IO-correspondence ↑		↑ *IO-correspondence*
/root/		/root + affix/

Benua, following the tradition of lexical phonology, proposes to categorize English affixes into two classes: those which mostly ignore the stress patterns of the root words (Class 1) and those which preserve them (Class 2):

(9) Types of English Affixes
Class 1: *-al*, *-ate*, *-ic*, *-ity*, *-ous*, *-in*, etc.
Class 2: *-able*, *-er*, *-ful*, *-ist*, *-ness*, *un-*, etc.

(10)a illustrates the output to output constraint for Class 1 affixes, and (10)b, that for Class 2 affixes. Both constraints state that the second output must be similar to the first one, but with

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

differing strength: the constraint imposed by Class 1 affixes ranks lower than that by Class 2 affixes.

(10) Two OO-correspondence Relations

a. Class 1: OO_1-Identity

órigin	→	oríginal
↑		↑
/origin/		/origin + al/

b. Class 2: OO_2-Identity

óbvious	→	óbviousness
↑		↑
/obvious/		/obvious + ness/

c. OO_2-Identity >> OO_1-Identity

The tableau representations are provided in (11) and (12). Note that Benua has to have two recursions in each of which the same set of constraints must be evaluated. In the *órigin/oríginal* case represented in (11), the stress pattern of *órigin* is determined in the first recursion. This pattern does not play a role in the determination of that of *oríginal* taking place in the second recursion because the weak constraint imposed by Class 1 affix *-al* (OO_1-Identity) is outranked by the regular stress-determining constraint (Align-R). In the *óbvious/óbviousness* case depicted in (12), the strong constraint imposed by the Class 2 affix *-ness* (OO_2-Identity) is ranked higher than the regular stress-determining constraint (Align-R), and thus the stress pattern of *óbviousness* determined in the second recursion is affected by the stress pattern of *óbvious* determined in the first recursion.

(11) Recursion (A)

/origin/	NONFINAL	ALGIN-R	OO_1-IDENTITY	>>
a. o(rí.gin)	*!			
b. ☞(ó.ri)gin		*		
c. (ó.ri)gin		*		

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Recursion (B)

>>	/origin + al/	NONFINAL	ALGIN-R	OO_1-IDENTITY
	a. o(rí.gi)nal		**	
	b. ☞o(rí.gi)nal		**	*
	c. (ó.ri)gin.nal		***!	

(12) Recursion (A)

/obvious/	NONFINAL	OO_2-IDENTITY	ALGIN-R	>>
a. ob(ví.ous)	*!			
b. (ób) vi.ous			**	
c. ☞ (ób)vi.ous			**	

Recursion (B)

>>	/obvious + ness/	NONFINAL	OO_2-IDENTITY	ALGIN-R
	a. ob(ví.ous)ness			**
	b. ob(ví.ous)ness		*!	**
	c. ☞ (ób)vi.ous.ness			**

There are some problems with this optimal theoretic analysis that should be pointed out. First, there is no explanation provided of the fact that there are complex words with a Class 1 suffix that retain the stress patterns of the root words (e.g. *óxygen-àte*). Second, the analysis makes use of an output-to-output constraint involving recursive application in a framework that supposedly dispenses with "derivation". Tableau in (11) undoubtedly involves a process metaphor, and these recursions are simply cycles in a modern disguise. Third, there is no straightforward way to bring semantic information to bear on a phonological process (although in Section IV.2 we demonstrated that the contrast between *demócratìze* and *géneralìze* can only be explained by the different degrees of semantic decomposability involved).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

VI. CONCLUSION

In this paper, we have shown that the theory of cognitive grammar, whose focus has primarily been on semantic analyses, is capable of offering a framework in which a phonological phenomenon such as interlexical relations in English stress can successfully be accounted for.[16] In Section III, we specifically observed that one can successfully account for interlexical relations by treating actually-occurring expressions as constraints and that cognitive grammar, being a usage-based, non-reductionist framework, can do so in a coherent manner by giving no inherent distinction between actually-occurring expressions and the schemas that are extracted from them. Optimality theory, on the other hand, can accommodate interlexical relations, but only with some fundamental theoretical incongruity, i.e. forcing an output to influence another output.

In Section IV.2, we further observed that if one wishes to offer a precise account of interlexical relations, one should employ semantic factors in addition to phonological ones. Optimality theory may well be able to incorporate semantic factors in phonological analyses, but only with significant conceptual or organizational difficulty. Cognitive grammar, in contrast, can naturally bring semantic factors to bear on phonological analyses because it employs only those theoretical constructs which have already been proposed elsewhere for semantic analyses. This demonstrates that unlike modular theories of grammar, cognitive grammar achieves theoretical unification, employing the same set of constructs to explain structures at both the phonological and semantic poles.

NOTES

1. We thank Matthew Chen and Ron Langacker, who gave valuable comments on earlier versions of this paper. We are also grateful to the editor of this volume and anonymous referees for their helpful comments. All the remaining errors are of course ours. The work reported in this paper was partially supported by the Keio Gijuku Academic Development Funds.

2. Words with certain affixes (e.g. *-able*, *-ful*, *-ness*) are always faithful to the stress patterns of the root words. In this paper, we will only be concerned with those affixes which would affect the stress patterns of at least some, if not all, root words.

3. For an illustration of various types of interlexical relations that affect the stress patterns of complex words, see Chen (1989).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

4. Therefore, cognitive grammar is in a sharp contrast with generative frameworks, which are characterized as "top-down, reductionist, and minimalist".

5. American pronunciation and American phonetic notation, not the IPA, will be used throughout this paper to follow the notation of Langacker (1988), to which the current paper owes much.

6. The semantic pole [THING] is to be taken as the semantic value common to all types of nouns, and the phonological pole [x] simply stands for any phonological content.

7. For discussions of the linguistic importance of the notion of prototype and its definition, the reader is referred to Lakoff (1987) and Geeraerts (1989, 1997), among others.

8. There are subschemas below the schema in question, but they are abbreviated for the sake of simplicity.

9. We believe that the range of data that are explainable by the two proposals are identical; one only needs to translate between selection among schemas and that among candidates. One can of course choose between the two models on the basis of "psychological reality"; however, the relevant mental activities that are involved are highly abstract, which leads us to believe that it is not possible to detect any decisive differences. We propose (5) here because it is more congruous with the interactive activation model and it is easier to mentally manipulate concrete entities (such as candidate expressions) than abstract entities (such as schemas).

10. In the figure, linguistic units listed in the grammar, i.e. conventionalized expressions, are enclosed in a rectangular box (the categorizing units in Figs. 4a, c, e, and n) and those not, i.e. unconventionalized expressions, listed in a box with round corners (the candidate expressions in Figs. 4y and z).

11. In the current paper, we adopt the "maximal onset" principle (Kahn 1976) for syllabification. There are other syllabification rules that treat "v" in *avocado* as the coda of the initial syllable or as "ambisyllabic", which can easily be accommodated in the current analysis only with minor representational modifications.

12. This symbolic relation always holds between the semantic and the phonological pole of any expression. However, these relationships were suppressed in Fig. 1 above for expository purposes, although they are actually present.

13. The exact nature of the negligible conflict will be described later in this section.

14. The details of this schema will be provided later in this section.

15. However, there have been some "noncyclical" approaches to English stress proposed within generative phonology. See Schane (1975, 1979), among others.

16. It would be due at this juncture to point out the limitations of the analysis presented in the current paper that should be addressed in future research. First, the range of data that are explained in the analysis is admittedly small, although we believe that they are representative and that the analysis can extend to a full range of data without significant modifications. Second, with respect to the activation principle in (5)a, objective methods to assess the degree of cognitive salience of a unit and the distance between a categorizing unit and a candidate are called for. Psycholinguistic experiments or simulation using a connectionist model could offer such methods.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

REFERENCES

Benua, L. (1997). *Transderivational identity: Phonological relations between words*. Doctoral dissertation. University of Massachusetts, Amherst.

Benua, L. (2000). *Relations between words*. New York: Garland.

Benua, L. (2004). Transderivational identity: Phonological relations between words. In J. McCarthy (Ed.), *Optimality theory in phonology: A reader*. Malden: Blackwell, pp. 419-437.

Chen, M. (1989). The English stress cycle and interlexical relations. Manuscript. University of California, San Diego.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Elman, J. L., & McClelland, J. L. (1984). Speech perception as a cognitive process: The interactive activation model. In N. Lass (Ed.), *Speech and language*, Volume 10. New York: Academic Press, pp. 337-374.

Farrell, P. (1990). Spanish stress: A cognitive analysis. *Hispanic Linguistics*, 4, 21-56.

Geeraerts, D. (1989). Prospects and problems of prototype theory. *Linguistics*, 27, 587-612.

Geeraerts, D. (1997). *Diachronic prototype semantics: A contribution to historical lexicology*. Oxford: Clarendon.

Halle, M., & Vergnaud, J.-R. (1987). *An essay on stress*. Cambridge: MIT Press.

Hayes, B. (1982). Extrametricality and English stress. *Linguistic Inquiry*, 13, 227-276.

Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry*, 15, 33-74.

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. Doctoral dissertation. Massachusetts Institute of Technology.

Kiparsky, P. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, 10, 421-442.

Kiparsky, P. (1982). From cyclic phonology to lexical phonology. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representation*, Part 1. Dordrecht: Foris, pp. 131-177.

Kumashiro, F. (2000). *Phonotactic interactions: A non-reductionist approach to phonology*. Doctoral dissertation. University of California, San Diego.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Kumashiro, T. (1990). The cognitive basis of grammar: A noncyclic analysis of English cyclic stress. Manuscript. University of California, San Diego.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: Chicago University Press.

Langacker, R. W. (1987). *Foundations of cognitive grammar,* Volume 1*, Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, R. W. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics*. Amsterdam: John Benjamins, pp. 127-161.

Langacker, R. W. (1990). *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Langacker, R. W. (1991). *Foundations of cognitive grammar,* Volume 2, *Descriptive application*. Stanford: Stanford University Press.

Langacker, R. W. (1999). *Grammar and conceptualization.* Berlin: Mouton de Gruyter.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249-336.

Prince, A. S. (1983). Relating to the grid. *Linguistic Inquiry*, 14, 19-100.

Rubba, J. E. (1993). *Discontinuous morphology in modern Aramaic.* Doctoral dissertation. University of California, San Diego.

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science*, 9, 75-112.

Schane, S. A. (1975). Noncyclic English word stress. In D. Goyvaerts & G. Pullum (Eds.), *Essays on the sound pattern of English*. Ghent: E. Story-Scientia, pp. 249-259.

Schane, S. A. (1979). The rhythmic nature of English word accentuation. *Language*, 55, 559-602.

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.

Waltz, D. L., & Pollack, J. B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, 9, 51-74.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Towards a Usage-Based Cognitive Phonology[1]

GITTE KRISTIANSEN[*]
Universidad Complutense de Madrid

ABSTRACT

The usage-based conception of language is a major tenet in Cognitive Linguistics, but cognitive phonology has not yet been developed sufficiently in this direction. Often, phonemic analysis is carried out at the high level of abstraction of `a language´, disregarding rich patterns of language-internal variation. This paper first argues that cognitive phonology must aim at a higher degree of descriptive refinement, especially in the direction of social variation. Then it goes on to examine the implications of a usage-based and multi-faceted model for a theoretical discussion of the phoneme as a prototype category.

KEYWORDS: usage-based cognitive phonology, lectal variation, distributed cognition, receptive and active competence, language acquisition, models of phonemic representation

[*] ***Address for correspondence***: Gitte Kristiansen. Universidad Complutense de Madrid. Facultad de Filología. Departamento de Filología Inglesa I. Lengua y Lingüística. 28040 Madrid. Phone: +34913945392. Fax: +34 913945478. e-mail: gkristia@filol.ucm.es

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I. ON THE NECESSITY OF A USAGE-BASED FRAMEWORK FOR COGNITIVE PHONOLOGY

A recent Google search on `accent reduction´ (January 7, 2007) gave 2.040.000 results. Most of the first 400 results corresponded to courses offered by private companies or public institutions (universities included) aiming at a reduction of foreign accents. However, a surprisingly high number also had as their target the reduction of native accents. In such cases, the course descriptions often explain in quite straightforward terms that having a regional accent can `keep you from being promoted´ or `hamper you professionally and socially´. It is easy to find companies that specialize in `American Regional Dialect Reduction´ and offer courses or products which teach the unfortunate speakers with regionalisms such as `American South and Texas´, `Mid-West farm Belt´, `American Urban and Rural Black´ or even `New York City and North Jersey´ how to `speak without an accent´. It has thus not gone unnoticed that there is market for teaching people how to speak with socially prestigious accents. The unspoken assumption is that accents are socially diagnostic: humans have `passive´ competence of lectal varieties (cf. Kristiansen, 2003) in the sense that we possess the ability to process clusters of linguistics cues which quickly and efficiently signal social and regional origin (hearer *categorizes* speaker lectally and socially) and concomitantly invoke the corresponding social stereotypes (hearer *characterizes* speaker socially and geographically).

One of the most important cues to correct dialect identification (perhaps even the most significant one: cf. Purnell *et al.*, 1999; van Bezooijen & Gooskens, 1999) is allophonic variation. However, while the disciplines of Sociolinguistics and Social Psychology of Language have explored the relationship between linguistic variants and social meaning for more than four decades now, social and regional variation has strangely enough been relegated to a secondary position in cognitive phonology. Yet, there are good reasons for investigating these areas in more depth within a Cognitive Linguistics framework. For a start, it would seem more than just pertinent, the emphasis on meaning-making processes provided, to investigate not only the mechanisms by means of which allophonic variation evokes social meaning, but also the various options language users employ once the link between linguistic form and social domain has been established (cf. Kristiansen forthcoming a). Further, an analysis of linguistic variants in actual

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

usage within social and linguistic dimensions which taxonomically speaking are more specific than the large-scale and abstract category of `a language´ is only in line with the fundamental claim that Cognitive Linguistics is a *usage-based* perspective. As Langacker, the scholar who coined the term, once phrased it:

> In a usage-based model, substantial importance is given to the actual use of the linguistic system and a speaker's knowledge of this use; […] It is a non-reductive approach to linguistic structure that employs fully articulated schematic networks and emphasizes the importance of low-level schemas (Langacker, 1999: 91).

In turn, Geeraerts (2001, 2005; Geeraerts *et al*., 1994) has repeatedly drawn attention towards the logical entailment of such a position: we can only take the claim that Cognitive Linguistics is a usage-based approach seriously if the kind of language that we analyze is *real* language, language as it is actually used by real speakers in real situations. Obviously, an analysis which in a natural way incorporates social factors and other types of language-internal dimensions not only provides us with a far more realistic picture, but also widens the gap with respect to disciplines according to which it is still acceptable, and possible, to analyze languages in terms of idealized speakers and homogeneous speech communities.[2] It naturally follows that a fine-grained map of allophonic variation within a given speech community cannot ignore social variation of the type just described.[3] Lectal variation will naturally form part of any description, or model, which purports to provide a realistic picture of phonetic variation and phonemic categorization.

However, the implications are not only descriptive, but also theoretical. From a diachronic perspective, knowledge about the social significance of linguistic variants may well turn out to have an influence on active competence, and accordingly on the nature of the variants in actual usage – both quantitatively and qualitatively - in a given language at a given historical time. A theory of language that intends to describe and explain the dynamics of language change in adequate ways cannot afford to ignore the (synchronic) mechanisms of actual language usage. Obviously, to obtain a picture which allows us to discuss the nature and evolution of not only phoneme categories but also phoneme inventories, the very first step is to aim at a high level of

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

descriptive accuracy. Usage-based models, then, take the data as they actually appear and set up theories which conform to the facts.

As far as the link between usage and language acquisition is concerned, Taylor (2002: 27) summarizes the relationship in the following way:

> It is assumed that the input to language acquisition are encounters with actual linguistic expressions, fully specified in their phonological, semantic, and symbolic aspects. Knowledge of a language is based in knowledge of actual usage and of generalizations made over usage events. Language acquisition is therefore a bottom-up process, driven by linguistic experience.

Though I shall also touch on acquisition in what follows, my main concern will be the descriptive and theoretical implications of zooming in on real usage in a more persistent way.

For the sake of highlighting the importance of social variation, I have started out with this particular aspect, but a usage-based model is obviously not one-dimensional. In fact, if Cognitive Linguistics shares an interest in low-level schemas situated at the level of parole with Sociolinguistics, with Functionalism it shares, amongst many other aspects (cf. Nuyts, 2005) the view that meaning cannot be studied in isolation, separable from the nature and the purpose of what is communicated and from the dynamics of communication as such. In the next section I focus on three of the dimensions which have a bearing on allophonic variation, viz. the ideational, discursive, and social functions of language. In other words, I will discuss (a) the kind of phonetic variation which serves the fundamental purpose of realizing a distinctive unit (i.e. the phoneme) and which does not necessarily carry an additional social message, (b) co-textual variation (that which derives from discursive factors), and (c) variation of a contextual type (that which pertains to social cognition and interaction).

II. A MULTI-DIMENSIONAL APPROACH TO PHONETIC VARIATION

In what follows I shall discuss a number of different dimensions implementing the terms `ideational´, `social´ and `discursive´. These correspond, roughly speaking, to Halliday's (1978) trichotomy of the `ideational´, `interpersonal´, and `textual´ functions of language. However, on

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

the assumption that not only inter-personal, but also inter-group and intra-group mechanisms are at work in social interaction, the term `social´ will be employed instead of `interpersonal´. Likewise, `discursive´ is preferred to `textual´, since, although we know that texts – in the sense of coherent stretches of language regardless of size or mode – are oral as well as written, the dynamics of ongoing, oral, discourse is made more prominent by using the former expression. I retain, albeit reluctantly, the term `ideational´ since social information is no less a question of (internally complex and subjectively construed) semantic domains than is `factual´ information. In this respect, Jakobson's (1960) `referential´ and Lyon's (1977) `descriptive´ do not offer a helpful distinction, either.

II.1. The ideational function of language

The ideational function of language, according to which phonemes function as builders of `factual´ meaning at the level of the morpheme and allophonic variants are processed as mere phonemic slot-fillers, is so well-known that little needs to be explicated. In very general terms, since the discovery of minimal pairs, both Structuralism and Functionalism established a clear distinction between the meaning-making levels of linguistic structure (beginning with morphology) and the meaning-building ones (i.e. the levels `below´ morphology). It was not until the birth of Sociolinguistics that the correlation between social variables and phonetic variants began to be studied in a systematic way. Cognitive Linguistics has perhaps neglected precisely that aspect, but in turn it has contributed to a prototype-theoretical conception of the phoneme as a mental category (Mompeán, 2004; Nathan, 1986, 1994, 1996, 1999; Taylor, 1990, 1995, 2002). In a phonemic prototype category, phonetic variants cluster around a central member which is maximally contrastive with respect to the central members of neighboring categories (Taylor, 1995: 221). In theory, then, the specific sets of phoneme categories organized around such major acoustic-perceptual contrasts in actual usage (the options logically being constrained by human physiological conditions in general) constitute the phoneme inventories of natural languages.

However, although the cognitive phonology view of phoneme category structure is certainly an attractive model, it is far from unproblematic. For a start, it is implicitly assumed that

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

the vast kind of language-internal variation we encounter across lectal categories can be brought under one schematic representation. It is assumed, for example, that the English phoneme category /t/ can be described (implementing either a network model or a radial category model) in such a way that all the variants of /t/ in actual usage in `English´ can be subsumed within the same category. In the case of /t/, even if social variants are allowed to become reflected, we can still easily talk about a consistent radial category, featuring extensions based on relative similarity with respect to central or more peripheral members, all of which ultimately organized, in more or less direct ways, around a prototypical member. Yet in many other cases, the model is not that easily applied. Or rather, viewed from a usage-based perspective, the data fail to fit the model in as neat a way as we would like them to. This is particularly true in the case of the vowels. Let us take an example from the Linguistic Atlas of England (Orton *et al.*, 1978):

Figure 1. Variants of the vowel in the verb *lay* in traditional English dialects according to *The Linguistic Atlas of England* (Orton *et al*., 1978).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The map in Figure 1 represents realizations of the vowel pronounced in the word-form <lay> in traditional dialects in England between 1950 and 1960. The data are perhaps somehow obsolete, but they will do for our present purposes, as the heterogeneous situation conveyed by the map is representative of dialectal variation. We may note, for a start, that many of the steps of the Great Vowel Shift never reached the very North of England, with retention of many of the monophthongs from the Old English period as a result. In fact, while it is certainly still possible to draw up a schematic representation which depicts chaining relationships among the variants, or instantiations, in actual occurrence, it is much harder to posit that there should be just one single phoneme (not in the sense of a prototype category, but of a distinctive builder of ideational meaning) at work at the same time. We encounter realizations such as [laɪ] and [li:], which would clearly be transphonemic if one specific phonological system, such as RP, were to form the basis of our model (i.e. [laɪ] and [li:] evoke the semantic poles of *lie* and *lee* for a speaker of RP and many Southern English dialects). The question is not only how to chart the internal structure of a formal category in adequate ways, but also whose system(s) we are representing. Large-scale speech communities are complex and heterogeneous, and to be really usage-based cognitive phonology must account for the fact that there are multiple, and quite dissimilar variants in use and different phonological systems at work within the same language. The easy way out is to adopt the position that each lectal variety forms an autonomous system of its own, to be analyzed independently, regardless of the existence of neighboring, or adjacent systems. As I have pointed out on other occasions (cf. Kristiansen, 2003: 76), this is precisely the way systemic-functional linguistics solved the problem. By establishing a distinction between language as institution (consisting of independently formed varieties) and language as system (language perceived as a system analyzable in terms of layers of linguistic structure), Halliday (1978) sifted one fairly homogeneous (but of course still very heterogeneous in many other respects) kind of language from the sum of its lectal varieties. Yet a usage-based approach cannot afford to work at such a high level of abstraction, especially in the field of cognitive phonology. It is not realistic to work around one variety of a language, no matter how prestigious and accessible it happens to be, and equate it with the language in question.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

II.2. The discursive function of language

Natural phonology argues that the discursive roles of hearer and speaker result in a series of (perhaps conflicting, but on the other hand perfectly compatible) tendencies. While both hearer and speaker seek improvements, speaker does so by means of mechanisms involving `ease of effort´ and hence `ease of articulation´. The hearer-oriented role, in turn, is rather aimed towards `clear and effective communication´. Following Stampe (1979), Nathan (1996: 116-117) has argued that *fortitions* are processes which select among all possible human sounds those which constitute the phonemes of a particular language and which define prototype effects for language sounds. *Lenitions*, on the other hand, are processes which create allophonic variants, i.e. extensions from a prototypical member within a radial network. These phonological processes are "universal cognitive mechanisms that languages have and may or may not use (that is, may or may not suppress) in any given instance" (Nathan, 1996: 113). Fortitions thus help us understand phoneme inventories in terms of series of phoneme categories whose prototypical members show a maximum degree of perceptual difference. Such an approach is of course highly compatible with general principles of prototypicality: maximum acoustic-perceptual salience and inter-categorial contrast as an organizing principle for the structure of phoneme inventories and categories constitute criterial factors in both cases. Numerous processes of assimilation, on the other hand, may be explained in terms of lenitions (e.g. the `rule´ according to which an alveolar nasal becomes labialized in immediate contact with other labial sounds: <in Paris> [ɪm'pærɪs]). Obviously, the discursive (or co-textual) variants which systematically occur in a given dialect - or at least in the most prototypical instantiations of the variety in question – would naturally form part of a minute description of intra-phonemic variation.[4]

II.3. The social function of language

Accents are socially diagnostic. When a stretch of speech is processed, not only can hearer decode an ideational message (conveyed by constructions from those layers of a language we traditionally label in terms of phonology, morphology, syntax or lexis), but also an additional social message. This can be even more forceful than an explicit statement, precisely because

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

hearer on occasions receives the information in an implicit way. The information we receive without being fully aware of it is likely not to be questioned in the same way as an explicit statement about regional origin and psychological characteristics would. Paradigmatic variation thus overrides the linear constraint of language in subtle ways. To the extent that children gradually build up knowledge of lectal varieties and learn how to relate them systematically to social domains, social values and stereotypical perceptions, a speaker can become not only categorized, but also characterized on the mere basis of his accent.

If this (relative) awareness has implications for language change, it may ultimately also have an influence on the shape of phonemic categories and inventories. In this respect, when a given phonetic variant begins to spread throughout the social and regional dimensions, ceases at a given point, co-exists with other variants and eventually replaces a number of its local `competitors´, the role and the motivations of the speakers who, either above or below the level of conscious awareness, opt for just this variant should not be underestimated. In fact, as Bybee points out, in order to understand language change, we should look for dynamic mechanisms which pertain to or influence actual language usage:

> …the true universals of language are the dynamic mechanisms that cause language to change in certain systematic ways as it is used and as it is transmitted to new generations. (Bybee, 2001: 189)

The kind of social cognition that will be described in more detail in this section can be viewed to fall within the category of such dynamic mechanisms. In a multi-dimensional approach to language variation and change, cognitive, social and functional factors will naturally combine as causative mechanisms to eventually shape the systems of a given language. This is a conception which is ultimately *not* compatible with theories which regard social factors as mere triggers of deeper, inherent or `natural´ causes of language change (e.g. Aitchison, 1991).

At this stage it is necessary to clarify that in this paper the term `social´ is used to denote the various social contexts which surround us. On the one hand there is an *immediate context* which involves the speech event as such, including the participants (speaker, hearer, bystanders, etc.) and their social status, the setting, the topics, the communicative goals, etc.). On the other

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

hand, there is also a *wider social context* which involves the participants' knowledge of, or belief in, a series of the Cultural Cognitive Models (Holland & Quinn 1987) at their disposal, and their knowledge about social groups, including social and linguistic stereotypes. These two types of context intertwine and interact in intricate ways. In fact, they are only distinguished in such an apparently easy and discrete way here for the sake of explanatory clarity. The former situates social interaction in real time and space and enables us to work around actual (and often purposive) usage of speech styles. The latter is transmitted, negotiated, developed and maintained in situated social interaction. Both types of context thus play a crucial role in social cognition.

In very general terms, in the author's main line of research (e.g. Kristiansen, 2001, 2003, forthcoming a) the fact that accents are socially diagnostic has served as a starting-point from which related issues have been explored in a number of different directions, including the relationship between accents (in terms of structured speech patterns) and social meaning from a Cognitive Linguistics perspective. My research thus falls within the wider fields of cognitive dialectology, cognitive sociolinguistics and language variation and change, but part of the analysis has a direct bearing on phonology and may be summarized as follows:

1. Lectal varieties (i.e. those categories we more traditionally label in terms of regional or social dialects, accents or speech styles) and social categories (i.e. social groups and identities such as British, Cockney, Northerner, South African, Australian, etc.) constitute prototype categories which interact at various levels of abstraction. The central images of lectal varieties (speech templates or linguistic stereo*types*, consisting of a cluster of salient features, of which allophonic variants play an important role) act as effective reference point constructions which through a basic metonymic operation (accents form part of a wider frame, a social domain) evokes the corresponding social stereotypes: LANGUAGE STANDS FOR SOCIAL IDENTITIES.

2. Once the link between a linguistic and a social stereotype has been established, it may be put to even more constructive uses, as speakers posses not only receptive, but also – at least to some extent – productive competence of speech styles. It is extremely difficult to imitate a

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

non-native speech style to perfection, but the most salient features are relatively easy to perform. Paradigmatic variation is also a metaphor: LANGUAGE IS A TOOL FOR CONVEYING SOCIAL MEANING.

These are statements which need to be spelled out in more detail. When I implement the notion of linguistic stereotype it is in a neutral, technical way. What I have in mind is a complex cluster of features which, from the perspective of folk perception, in the best and clearest way allows us to categorize and identify the structured speech style of the members of a given speech community. Linguistic stereotypes very effectively evoke the corresponding social stereotypes, conceived, in equally neutral terms, as outgroup images of a given social category.

Speech patterns acquired in early childhood are not easily changed, and linguistic stereotypes thus constitute an especially reliable marker of social identity. Hence, the link between linguistic and social stereotypes is fundamentally of a metonymic nature: an EFFECT FOR CAUSE mapping which leads the conceptualizer from a linguistic trigger to a wider social target: to a social group and the encyclopaedic knowledge we have about it (social habits, dress, dance, song). This knowledge includes a series of Cultural Cognitive Models (i.e. ideological patterns and components) and often take the form of stereotypical perceptions.

I understand social stereotypes in terms of simplified outgroup perceptions which condense information regarding what the members of a given group are like (e.g. in terms of psychological attributes, ideological beliefs and social behavior). The existence of a fairly stable relationship between speech styles and social targets thus underlies the general metonymy LANGUAGE STANDS FOR SOCIAL IDENTITIES. When a linguistic feature is heard as `prestigious´, `intelligent´ or `posh´ and in reality it is the group of speakers associated with the feature in question which is being evaluated as such, it is accordingly not a process of iconization (Irvine & Gal, 2000), but rather an indexical process which is at work. The force of the process is easily comprehended if we bear in mind that the features contained in just one two-syllable word (cf. Purnell *et al*., 1999) suffice to evoke the whole lectal category (a part-whole metonymy) which in turn links, again indexically, with a social domain and the corresponding social stereotypes.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Receptive competence is however only the starting-point of a much more complex story. Once a stable referential link has been established between a lectal variety and a social domain, entrenched linguistic trigger - social value relationships can now be put to new, constructive uses. In other words, the existence of receptive competence can be exploited by speakers in order to *signal* social values. Structured paradigmatic variation can thus also be put to metaphorical uses: LANGUAGE IS A TOOL FOR CONVEYING SOCIAL MEANING.

If I speak of the central images of lectal varieties and social groups as cognitive reference point constructions (and not just cognitive reference points), it is to emphasize the fact that such images are relative and relational: situated and group-dependent construals. Social and linguistic stereotypes are perhaps best understood as instances of situated cognition (cf. Kristiansen, forthcoming b), as structure which emerges, is transmitted, negotiated and maintained through dynamic interaction between people in real situations in real historical time. In consequence, the specific combination of items which compose a given social stereotype will to a large extent depend on contextually determined intergroup relationships. In the case of linguistic stereotypes, the features that characterize the speech of an outgroup will also be determined by the nature of the features present in our own ingroup speech pattern. For instance, the phonetic variants which are perceived as salient and identifying of French to an Englishman or a Dane might not be viewed as such by an Italian. If a speaker possesses a similar realization in similar phonetic contexts in his own language, the feature will not stand out as perceptually salient. Lectal categorization, when viewed from this broad perspective, is not only a purely cognitive phenomenon, but also a process which must also be considered in terms of cultural situatedness.

For linguistic stereotypes to relate in an exclusive way to a social group – for accents to be socially *diagnostic* – the cluster of features that set a speech style off as distinct from others must be composed of a unique combination of perceptually salient features. This is in consonance with Nunberg's (1978) line of thought when he asked himself how hearer and speaker manage to determine referents in deferred ostension. In metonymic conceptual operations, what kind of visual demonstratum will successfully lead to the intended referent, and which will fail to do so? As I have previously pointed out (Kristiansen, 2003), Nunberg in reality asked himself what the ideal signifier in operations of metonymic reference is like. He reasoned that a given form will

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

successfully lead to the intended referent if it relates to it in an `exclusive´ way (so as to identify and not just characterize the target) and that the form in question must be `perceptually distinct´ enough for effective subclassification to take place, so as to be able to distinguish it from similar forms within the same general category which have different values attached to them.

In the field of phonology, there are multiple possibilities (though each language only exploits relatively few of these) of establishing acoustic-perceptual contrasts which allow for salient subsets to become effected within the more general category of a phoneme. There is no reason why such minor contrasts should not function according to the same fundamental principles as those which determine degrees of prototypicality for phoneme categories in general:

> …the putative central member of /t/ -say, the voiceless aspirated alveolar plosive- enters into a number of highly salient perceptual and articulatory contrasts with the putative central members of neighbouring categories, such as the unaspirated alveolar plosive of /d/, the voiceless aspirated velar plosive of /k/, and so on. (Taylor, 1995: 228)

Intraphonemic acoustic-perceptual contrasts (subphonemic prototypes) are thus minor when compared to that which sets [t^h] off from, say, [k^h], but still major enough to create distinct subsets *within* a given category:

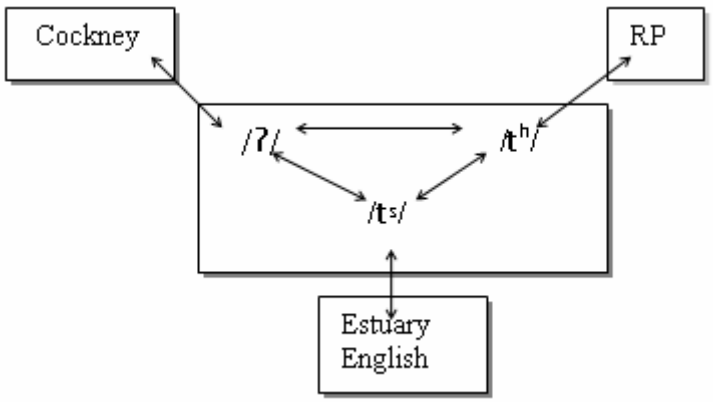


Figure 2. Some linguistic stereotypes based on intraphonemic acoustic-perceptual contrasts within the category /t/ in British English and metonymic reference to social domains.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

As a case in point, the use of the glottal stop in intervocalic position is obviously not sufficient to invoke the Cockney accent. Rather, it is a complex cluster of features (cf. Kristiansen, 2003, forthcoming a) which from a hearer-oriented perspective effectively categorizes a stretch of speech as a token of a given type in an exclusive way. From a speaker-oriented perspective, we would speak in terms of social differentiation being achieved by means of linguistic distinctiveness (cf. Giles *et al*., 1987; Tajfel & Turner, 1979).

Assimilation to a prototype category is usually thought of in terms of relative similarity with respect to another member of the category in question, be this peripheral or more central. Hence we speak of chaining relationships or radial networks. Prototype categories are flexible entities in the sense that the boundaries are extendable (new members may be added to allow for human cognition to adapt itself to change, innovation or new discoveries in a complex social and physical world). In theory, a new member is similar enough to at least one existing member in order to be categorized as a member of a given category and not as a member of a contrasting category (or an instance of a given schema, not another). Conversely, however, a new member must also necessarily be perceived as distinct, or different enough to deserve the status of a different subcategorization, a new extension or a new instantiation, to use several of the notions in current usage, and not just a token of an already existing type.

For the sake of exemplification, consider an invented case of categorization from the visual domain. In Figure 3, the left-most shape is a *kiki* and the right-hand shape is a *booba*.[5] Or so at least 95 per cent of subjects systematically estimated in a series of experiments on psycho-acoustics (Köhler, 1929, 1947; Werner, 1934, 1957; Werner & Wapner, 1952) when asked which shape was called what in a language unknown to them, the options being kiki and booba (the latter word pronounced with [ɔː], not [uː]):

Figure 3. A kiki and a booba.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Suppose that, as a reader of this paper, you are now asked to draw another kiki and that you draw a shape which is very similar to the kiki represented above. Suppose that you are told that you were expected to reflect a `different´ kiki, not the same kind of kiki (i.e. not a token of the same type). You then draw a kiki with fewer or perhaps more sides than the first one, but presumably still with pointed sides, so as not to produce a booba. If asked to do the same with the booba shape, your drawing might resemble this one:

Figure 4. A different kiki and a different booba.

Say more instances of irregular kikis, or six-sided kikis, become produced, either in different situations or by different groups of subjects. And that four-sided boobas show up systematically alongside seven-sided boobas under a series of contextual circumstances. The general categories of kikis and boobas have now become divided into subcategories, with slightly different tokens representing two different subtypes. At the linguistic end of the formal trigger-conceptual content axis, a new term would in all probability now arise to designate the new intra-categorical subtype, and this form would in many cases iconically convey both membership (retain part of the word-form designating the general category) and subclassification (possess some kind of formally distinctive element). The adjectival modifier in `four-sided kiki´ serves the latter purpose, the modified head the former.[6] The point is of course that successful subcategorization seems to be based as much on subtle, but still perceptually salient enough differences as on perceived similarity. In a similar way certain allophones, those which I have referred to in terms of subphonemic prototypes, while perceived as members of the same general category, are also perceptually distinct enough to form a new subcategory and thus serve as ideal triggers of new, additional meaning. Salient allophones can serve the dual purpose of realizing a phoneme (according to the ideational function) and evoke social group membership at the same time.[7] When viewed from both a hearer and speaker-oriented perspective, the possibility of drawing on such a pool of triggers of social meaning enables hearer to decode social information on the one

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

hand, and allows groups of speakers to encode it by choosing especially contrastive forms in their speech, much in the same way as they would, more mundanely, wear a distinctive kind of garment or opt for a specific hairstyle.

III. LECTAL VARIETIES AS EXPERIENTIALLY GROUNDED CONSTRUALS

In this section lectal categories will be examined in terms of construals grounded in individual and group-related experience. In III.1, I discuss the distinction between receptive and productive competence of lectal categories and relate competence to the notion of relative awareness. Subsection III.2 centers on lectal competence and language acquisition, and finally, in III.3 I address the question of lectal competence and distributed cognition.

III.1. Receptive and productive competence

Languages are schematic with respect to their instantiations: we inevitably speak a given variety of our mother tongue. In much the same way, linguistic input necessarily consists of real instantiations which become processed for a variety of purposes in terms of low-level or high-level schemas. For example, a phoneme is a schematic abstraction which cannot be pronounced as such. Also, words are invariably realized by means of a combination of specific phonetic variants. A word such as <butter> can only be realized as e.g. [ˈbʌtʰə], [ˈbʊtʰə] or [ˈbʌʔə] and then processed `ideationally´ as a sequence of phonemes (/ˈbʌtə/ → `butter´). In the right circumstances (i.e. when hearer possesses the necessary knowledge and is attentive enough) such instantiations will also be processed lectally (a given combination of linguistic features leads us to a particular lectal variety). That phonetic detail should be stored and processed alongside the function of realizing a phoneme is certainly not at odds with Bybee's (1988, 2001) model of a mental lexicon and a usage-based phonology; words are stored in their concrete phonetic forms and phonetic detail retained in long-term memory.

Let us assume that we gradually acquire knowledge about a large number of lectal varieties and the speech communities they relate to. In other words, that we gradually acquire receptive

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

competence of speech styles. Lectal categorization would then involve a conceptualizer who correlates a token (stretch of unidentified speech) with a number of idealized speech models (linguistic stereotypes). The similarity may be relative, of course; two people who `speak with the same accent´ obviously do not speak exactly the same way. Rather, their intonation patterns, phonetic realizations and phonemic slots are judged to be relatively similar when compared to a model. But we also soon learn how to put linguistic stereotypes to other, equally constructive uses. The fact that effective categorization seems to be based on a reduced series of highly salient features facilitates the process known as style-shifting. While it is extremely difficult, if not impossible, to imitate a non-native accent to perfection, with all the subtle phonotactic and distributional combinations of salient and much less salient variants, the components of a linguistic stereotype are fairly easy to imitate. The term productive competence denotes the use of features from a style which does not form part of a speaker's habitual repertoire.

A note is now necessary on awareness. Accents are socially diagnostic because linguistic cues index social meaning, or – in more technical terms – because a source-in-target metonymic conceptual operation mediates between structured sets of linguistic triggers and the social domain they project. However, this process is presumably often a below-the-level-of-consciousness affair. A little more than four decades ago, with the birth of Sociolinguistics, it became clear to many scholars that a systematic study of social dialects could not rely on the same methods as those traditionally implemented in the study of regional dialects. Eliciting informants' intuitions in a direct way, e.g. by means of the questionnaires used in many surveys, proved to be an inadequate procedure for a systematic description of social dialects for two major and interrelated reasons. On the one hand, social speech styles relate to contextual factors. Speakers vary their style according to situational factors such as setting and topic, or the style and status of the interviewer.[8] They also evaluate, consciously or not, the relative position of their own style on a social hierarchy of varieties according to variables such as prestige and stigmatization – and might accordingly over-represent the actual occurrence of features which rate high on the scale of prestige. On the other hand, actual usage, as the sociolinguists soon recognized, is often situated below the level of conscious awareness. In consequence, speakers' own perception of their speech style is likely not to coincide with actual usage, self-perception being potentially

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

distorted, inaccurate or – if we take the principle of the cognitive unconscious seriously – quite simply not fully accessible.[9] Awareness is furthermore a gradable dimension, and speakers are presumably often only conscious to a certain degree of the messages they receive when listening to stretches of speech. However, messages received below the level of conscious awareness are still received – and may lead to positive or negative evaluations, to rejection, admiration or imitation, if only on an apparently intuitive basis. Finally, and with regards to productive competence, that a speaker should occasionally imitate – or try to adopt on a regular basis – a feature from a given speech style because it is `fashionable´ is a commonly quoted explanation of sound change. Yet fashion can surely also be viewed as a variable which is ultimately dependent on hearer's perception of the hierarchical position of the lectal and social domains projected by a given linguistic feature. This perception invariably involves at least an implicit degree of awareness regarding the link between linguistic feature, social group and social meaning. Fashion could thus also be viewed as a cover term for a limited, but still productive degree of awareness as regards the processes which relate linguistic form to social meaning.

We all soon acquire a natural stylistic repertoire and moreover posses the ability to imitate new speech styles – or at least the most salient features of the styles of other groups. We may even set up new, local identities. These are often effected by selecting a series of socially meaningful features stemming from stable, large-scale social categories. Eckert (2004), for instance, reports on the various ways in which one particular British English feature (final /t/ with an audible release of aspiration), associated with the British as superior, intelligent and educated, has been put to a variety of different uses by American speakers of English (cf. Kristiansen, forthcoming a).[10]

The features that are imitated, or adopted and put to new uses, in the first place are those that stand out as especially salient. In this respect, not all contrasts are equally important in terms of acoustic-perceptual perception, and this is one more reason why we should treat the social function as separate dimension. It could be argued, for instance, that an account of phonetic variation which incorporates the variants in actual usage within the lectal varieties that compose a given language (alongside those that arise from discursive factors and other relevant functions) will already include the kind of social variation which is under scrutiny in this section, i.e. that by

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

lowering the level of abstraction so as to work around language-internal varieties and not `a language´, the social function is already duly covered. But if we did that, we would on the one hand remain at a descriptive level and fail to appreciate numerous factors which might well have a bearing on language change, and on the other hand miss out on the opportunity to investigate such factors in more depth. Language acquisition, for instance, is one of the areas in which it might be particularly fruitful to invest.

III.2. Lectal competence and language acquisition

So far research on accent-based speaker identification (e.g. Purnell *et al*., 1999; van Bezooijen & Gooskens, 1999; the many studies on speech identification in relation to Artificial Intelligence) and perceptual dialectology (e.g. Niedzielski & Preston, 2000) has largely concentrated on adult informants. Yet, if we adopt a usage-based approach and assume that the acquisition of speech styles is experientially grounded and that phonetic detail is stored as such (Bybee, 2001) and put to constructive uses, then there is a lack of empirical studies on the acquisition of receptive and productive competence of lectal varieties in children. When do children begin to construe low-level schemas, paying attention not only to what is said, but also to how it is said? If schemas are usage-based, at what age does dialect identification emerge, and how specific is it at different intervals of age? There is a need in other words to investigate, amongst others, the following factors:

- The degree to which children acquire receptive competence of accents at different intervals of age.
- The relative precision with which accents are identified at different intervals of age.
- The relative degree of awareness regarding the specific features which allow children to proceed to correct dialect identification.
- The relative capacity of children to imitate accents (productive competence) at different intervals of age.
- The factors which, age apart, have a bearing on lectal acquisition.
- The relative degree of awareness of children regarding the relationship between linguistic and social stereotypes.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

III.3. Lectal competence and distributed cognition

If "input to language acquisition are encounters with actual linguistic expressions, fully specified in their phonological, semantic and symbolic aspects" (Taylor, 2002: 27) and "knowledge of a language is based in knowledge of actual usage and of generalizations made over usage events" (ibid), it follows that not all individuals will possess the same kind of knowledge, nor always effect the same kind of generalizations. Or as Geeraerts (1997: 110) formulates it, a distinction between "a cumulative, macro-level picture of the language and the individual language user's micro-level knowledge" is most convenient:

> […] we would probably not want to maintain that all mature speakers of the language actively command the entire range of semasiological possibilities that are combined in the prototype-based descriptions. An alternative way to interpret diagrams […] is to think of them as representing the summed knowledge of language users at a certain moment in the development of the language (and then also, of course, the knowledge of an ideal language user). (ibid.)

From the perspective of a level of granularity which lies above that of the individual, Sharifian (2003) argues that the elements of cultural schemas are not shared by all members of a cultural network, but rather distributed across the minds. It is not by virtue of the belief in only one schema that one becomes a member of a cultural group, but the overall degree of how much a person draws on various cultural schemas that makes an individual a more or less representative member. Cultural schemas, or cultural cognitive models, thus thrive within groups and the group emerges as such, shaped and brought into existence by relatively shared beliefs, values and norms. In similar ways, the group also determines and is determined by relatively shared speech patterns such as `dialects´, `accents´ and `styles´ and by relatively shared social stereotypes.

In this respect it would be interesting to know more about how uniform folk perception of the way in which another social group speaks is across the members of a given ingroup. We might also want to know more about the extent to which linguistic stereotypes constitute relative construals across different cultural and lectal communities. To what extent, for instance, does the linguistic stereotype of Spanish differ when acquired by a Frenchman, an Italian and an

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Englishman, respectively? How do the features of one's own mother tongue – and motheraccent-influence our perception of what stands out as contrastive or salient? But even more important are the theoretical implications of viewing competence of a given language in terms of relatively shared and distributed knowledge.

IV. MODELS OF PHONEMIC CATEGORY STRUCTURE

The two prevailing models of phonemic category structure in Cognitive Linguistics (Mompeán, 2004: 436-444) are the *radial category model* and the *network model*. In this section we shall briefly discuss these in relation to a usage-based cognitive phonology.

The radial category model assumes that less prototypical members are organized around a prototypical member in terms of extensions assimilated to the category on the principle of relative similarity:

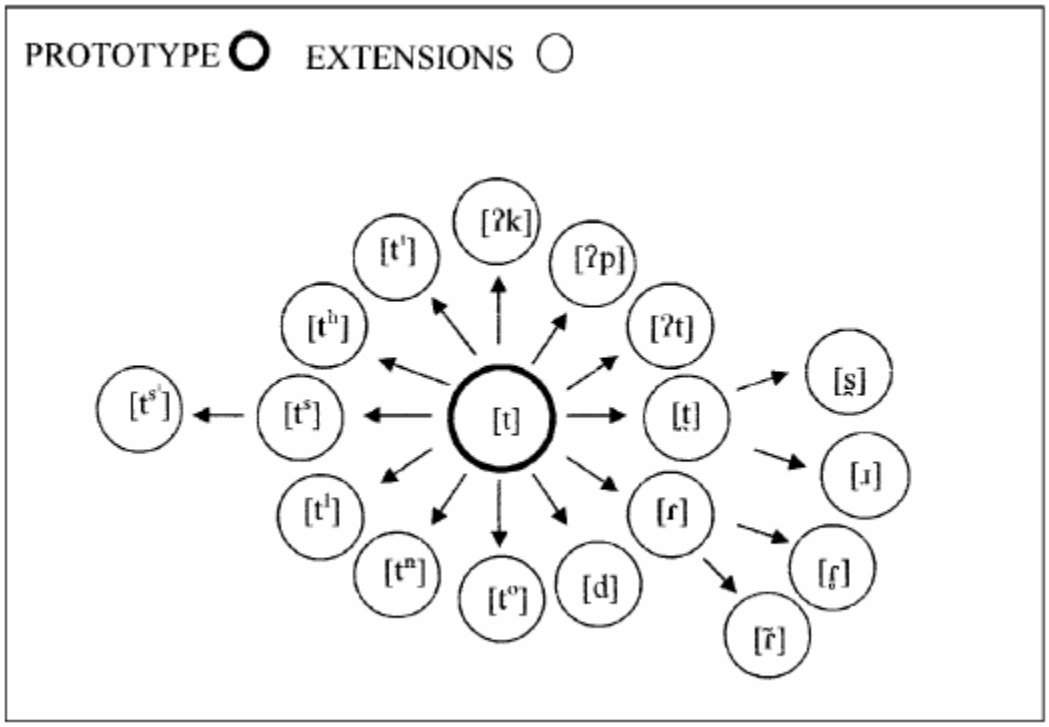


Figure 5. The radial category model as represented in Mompeán (2004).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The resulting chaining relationships stretching out from the centre to the periphery have been compared to the spokes of a wheel:

> The specific nature of the organization has been termed a radial category, because the relationship among the members is similar to an image of spokes on a wheel. There is (or may be) a central member or members. Arranged around the central members are less central ones, which are similar to the central member, but differ from it in some respect (Nathan, 1996: 110-111)

The members `arranged together on one spoke´ do not necessarily have any kind of relationship with members forming part of adjacent spokes:

> Adjacent spokes do not necessarily have any relationship with one another, but only via a path that they can both trace back to the same center. (Nathan, 1996: 112)

Metaphorical mappings involving source domains which comprise elements such as wheels and chains are adequate enough if lectal varieties are understood in terms of different analogical systems which do not interact in dynamic ways with one another. However, these mappings do not suffice if we also intend to describe the ways in which perceptual dissimilarity and intra-phonemic contrasts among the members of phonemic categories allow for language users to convey social differentiation through linguistic distinctiveness. Our pictorial representations, arrows included, should also convey the possibility that contrastive relationships emerge in a non-linear fashion. Metaphorically, then, we are in need of a more suitable source domain than wheel. The label radial network itself is in fact more neutral in the sense that it implies a system of interwoven relationships which need not follow a specific path.

Langacker's (1988) network model, on the other hand, involves a category prototype, context-induced extensions from this prototype and a schema which captures the commonality perceived in the various extensions. The improvement with respect to the radial category model thus lies in the fact that the network models operates with a two-level structure: the level of actual occurrence (of often quite dissimilar variants) and a schematic unit which is an abstraction over

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

usage-based events. In the case of phonemic categories, the schema corresponds to the phoneme as such:

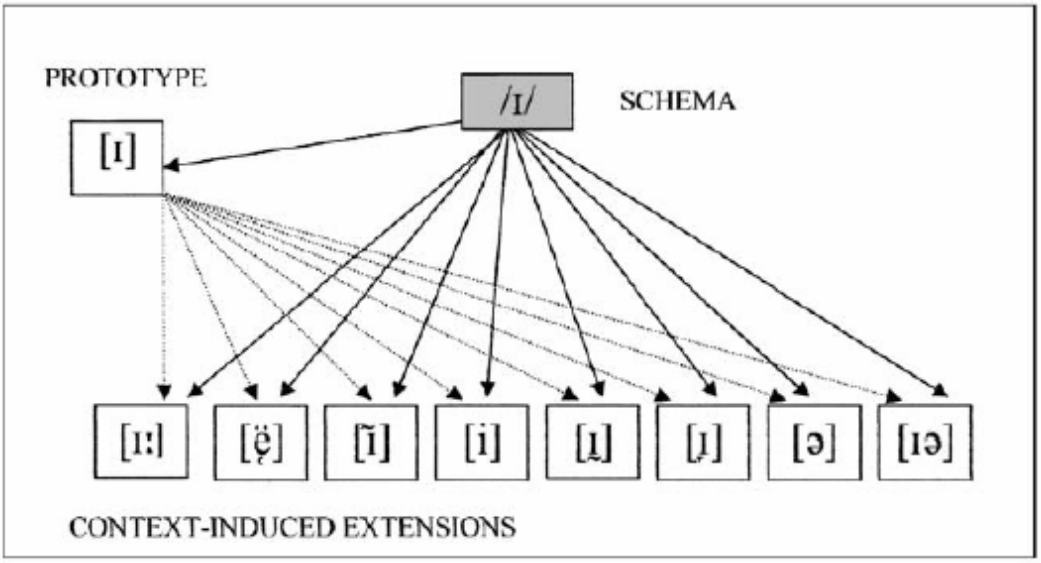


Figure 6. The network model as represented in Mompeán (2004).

As Mompeán explains, it is often not possible to extract one schema which represents a generalization with respect to all the extensions (which is only natural, the very nature of a prototype category considered). Rather, several schemas may arise which relate to different clusters of instances:

> However, it is not always possible to abstract a viable, psycholinguistically plausible schema that is fully compatible with all the members of a category. For example, not every member of the phoneme category /t/ shares the features ‘‘alveolar’’, ‘‘voiceless’’, and ‘‘stop’’, so the abstraction of a highly abstract schema which contains a feature common to all members of the category and distinguishes the category from others is impossible (Taylor, 1990). The model permits, however, the abstractions of local schemas embodying the commonality of many but not all members of the category (Bybee, 1999). Some commonality between certain members of a phoneme category may exist but the commonality may not extend to the totality of the members. One such local schema for /t/ could contain the features [voiceless], [alveolar], and [stop], shared by many but not all members of the category (Taylor, 1990). (Mompeán, 2004: 458)

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The need to posit that various schemas are at work at the same time is of course related to the fact that quite often formally very dissimilar variants are functionally operative as distinctive speech sounds in the same lexical sets in the same language. In such cases it is not easy to draw a coherent representation which involves only two levels of abstraction: specific language-internal variants and one language-specific schema.

In the light of the previous discussion, it might be useful to approach the problem from a variety of different angles. First, by applying our general knowledge about prototypicality to phonemic categories, second, by discussing the role of lectal varieties in phonemic description and third, by bringing in the perspectives of distributed cognition, expert analysis and folk perception.

IV.1. Phonemic categories and prototypicality

As we have just observed, similarity in form and commonality are considered as unifying factors in cognitive phonology. In the case of the radial network model, the relationship between the prototype and its extensions is based on perceived similarity, and in the case of the network model, schemas arise as generalizations embodying the commonality of their instances. In the absence of a commonality which applies to all members, local schemas obviously provide a category which exhibits a high degree of variation with internal cohesion. Let us observe that in a prototype category it is only normal for family resemblances to cluster in partially overlapping subsets, and we would certainly not expect there to be one single feature which would be common to all the extensions (unless one opts for a classical model based on necessary and sufficient features, or an essentialist definition). In fact, one would expect a phonemic category to exhibit the same characteristics as any other prototype category, to varying degrees, does: (i) absence of classical definitions, (ii) clustering of overlapping senses – or features, (iii) degrees of representativity, and (iv) absence of clear boundaries. In the case of lexical items (cf. Geeraerts *et al.*, 1994: 48), the first two characteristics operate at the intensional level and the last two at the extensional one, but combine in other ways as well: while (i) and (iv) reflect the flexibility and vagueness that characterizes many prototype categories, (ii) and (iii) in turn result from perceived

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

differences in structural weight. Nonequality and nonrigidity thus lie in the very nature of prototypicality.

Furthermore, it is often not formal, but functional criteria which determine whether a new subcategorization is established within a prototype category. In the case of the superordinate category FURNITURE, for instance, ashtrays may become assimilated as (peripheral) category members (Taylor, 1995) because they are functionally related to other such domestic artifacts despite being dissimilar in form to other central or non-central members. In the case of phonetic variants, from a speaker-oriented perspective the need for distinctive variants to convey social meaning can lead to both intra-phonemic and trans-phonemic variation. From a hearer-oriented perspective, dissimilarity would at first sight appear to be as counter-productive as lenitions are when viewed against the principles which underlie fortitions. However, opposing tendencies need not be incompatible. Even if a given speech sound (*y*) were to be classified as an instance of a given phoneme (*Y*) in hearer's own phonological system, when uttered by a speaker from a different speech community in a context where *x* is expected, *y* could still be classified as an instance of phoneme *X*, at least for the purpose of mutual understanding – as when [laɪ] is heard and `lay´ is understood (cf. above). In other words, if "each encounter with the language leaves a mental trace in the corpus" (Taylor, 2002: 33), and our receptive competence is experientially grounded, the variant y will be understood to belong to category X in the speech and system of the speaker in question – an at least ad hoc categorization based on functional, rather than perceptual criteria.

What, then, holds a phoneme category together? Clusters based on perceptual similarity or co-occurrence of distinct realizations in the same phonetic context, or more specifically, within the same lexical set? Is the train of thought primarily in the direction of `as this sound is similar to sound y it must be processed as phoneme Y´ or `as this sound occurs in a context where I expect sound y, it should be processed as an instantiation of phoneme Y´? In the former case, the radial category model is adequate enough, but in the latter case the network model is superior as it enables us to work around different layers of abstraction. It is also worth noticing that one of the implications of such a perspective is that phoneme recognition might well be lexically

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

mediated. If this is the case, the ideational function interacts with the social function of language in ways which are flexible enough to allow for both of these apparently conflicting tendencies to co-exist, to render language an efficient tool for a variety of communicative purposes at the same time.

IV.2. Phonemic categories and lectal variation

The network model, then, allows us to work around different levels of abstraction. Would it be possible for the model also to incorporate low-level schemas at an intermediate level of linguistic diversity? A model capable of operating with high-level schemas, local schemas and instances should indeed be able to reflect the taxonomic intricacies of language-internal variation in a far more precise fashion than the radial network model. It might also be fruitful to shift the perspective from category-internal variation (involving pure form) to a language-internal one, reflecting form, users and functions alike. An intermediate level of `local schemas´ would furthermore fill an important gap in a usage-based approach: that which – mediating between *parole* and *langue* – incorporates structured variation at the level of lects. Obviously, the kind of low-level schemas discussed in this section differ from the results of a post-hoc analysis concerned with finding patterns in subgroups with relative similarity as the basic criterial factor. Rather, it is an account which allows for language-internal categories to form part of the global picture. Relative dissimilarity between (clusters of relatively similar) features and a consideration of the ways in which these relate to their users would certainly also form part of the analysis.

The network model moreover allows us to reflect the factor of awareness as discussed in section III.1 above. The following three-level figure draws on Tuggy's (1993) representation of the ambiguity-vagueness cline between lexical polysemy and homophony:

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

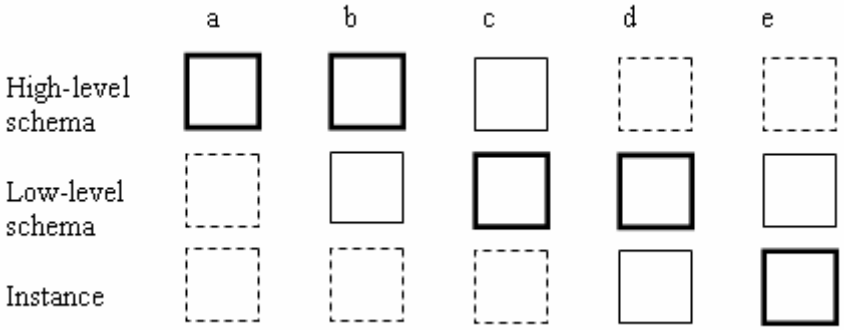


Figure 7. Salience and awareness in phonemic and lectal categorization

In Tuggy's analysis, the thickness and continuity of the lines iconically convey the enhanced entrenchment, or degree of salience, of either a schema or a semantic structure, associated with the same phonological pole. Tuggy views entrenchment in terms of enduring salience, i.e. salience apart from relatively transitory effects such as directed attention or heightened activation due to contextual factors. As it is precisely such transitory effects that we are interested in, the same conventions will do for our present purposes. 7a illustrates the ideational function of language: the high-level schema (the phoneme as a distinctive unit) receives full attention while lectal schemas and phonetic instances are backgrounded - which does not equal saying that the information provided is discarded.[11] Priority is given to the distinctive function fulfilled by the phoneme – an abstraction realized by specific instances. In 7b and 7c, lectal categorization becomes increasingly more prominent. 7d and 7e both illustrate a high degree of awareness regarding the social function of language: the high-level schema is backgrounded and attention is on the link between linguistic form and social meaning, with an emphasis on either form, or meaning, respectively.

IV.3. Phonemic representation and phonemic categorization: expert analysis vs. folk perception

We have so far argued that regardless of whether we implement the radial category model or the network model, our description should aim at being as refined as possible. Ideally, we would aim

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

at incorporating the whole range of variants in actual use in a given language - conceived of in terms of a heterogeneous speech community and a complex social system.

If what we intend, however, is to reach a better understanding of what goes on in the mind of the native speaker, this picture may well turn out to be rather fictitious. Even if we succeeded in bringing all the variants in actual use under the same schematic representation, the representation might not be very realistic. To the extent that cognition is distributed across speech communities and cultural groups - if knowledge is only relatively shared - a realistic view cannot posit that the individual user stores the whole range of variants in actual use within a language, be they phonetic variants or perhaps the multiple senses of a polysemous lexeme or preposition.

Sandra and Rice (1995) convincingly question the validity of representations of vast, global networks of senses depicted in the analysis of linguists as opposed to what ipso facto is acquired and stored in terms of mental representations in the mind of the individual, and it is indeed important to establish a distinction between the linguist's attempt at providing a global picture (one which cumulatively depicts the existence of multiple networks) and folk perception (the relative knowledge of a more global network in the mind of the individual - and the linguist's attempt at reflecting more local networks). The tension involves a clash between a perspective which in a structuralist or generativist fashion zooms in on `language structure´, independently of the fact that there might be more than one system at work, and one which regards language-internal variation as natural and worthy of attention. The two perspectives are not mutually exclusive, though. Both analyses are possible and complementary - but we need to acknowledge the differences in a clear and conscious manner.

V. CONCLUSIONS

Phonetic dissimilarity and categorization have been keywords throughout the various sections. Categorization as such is of course based on relative similarity - a cohesive factor which helps us organize a vast amount of variation into structured sets of like components - but it also involves the creation of subsets, established as much on the basis of relative dissimilarity. Categorization

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

also applies to the social world. At more precise levels of abstractions than that of `a language´, there are multiple social and lectal subsystems which together constitute `a society´ and `a language´. In this respect I have distinguished between the role of hearer and speaker, and argued that our receptive and productive competence of lectal varieties also plays a role in the configuration of phoneme categories and inventories. In other words, a cognitive dialectology - including a cognitive phonology - may well serve not only to mediate between `language´ and `society´ but also to spell out in full the consequences of a truly multi-faceted approach to phonetic variation.

I have also stressed the difference in perspective between expert analysis and folk perception in phonemic description, and argued that the distinction is useful when theoretical models face linguistic facts. Finally, I have examined the relative adequacy of the radial category model and the network model and concluded that the network model seems to present a number of advantages over the radial category model.

NOTES

1. This paper is associated with the research project HUM2005-08221-CO2-01.

2. E.g. the generativist and structuralist conceptions of language as a system (competence and langue, respectively) which is analysable independently of social and contextual factors.

3. For the importance of distributed cognition and the distinction between expert analysis and folk perception, cf. section IV.3.

4. Note that I intentionally use the term `dialect´, and not `language´: the specific nature of such low-level schemas varies considerably from one language-internal variety to another. This is one of the reasons why it is impossible for an average (adult) speaker to imitate an accent to perfection.

5. I am grateful to Raphael Berthele for introducing me to the `kiki´ and the `booba´ in his intervention on folk perception and phonosymbolism in the Theme Session *Lectal Variation and the Categorization of Lectal Varieties in Cognitive Linguistics*, ICLC9, Seoul. In the original experiment, Köhler (1929) called the stimuli `takete´ and `baluma´.

6. Ablaut of course serves the same differentiating purpose. In the case of English <swim, swam, swum>, the combination of three maximally distinct variants (close front i, open a, close back u) renders paradigmatic variation less ambiguous. The use of specific combinations of phonemes or morphemes in processes of derivation and declension is in this sense not entirely unmotivated. An interesting case is that of the terms <starboard and <larboard>. Both word-forms used to denote the left and right side of a ship, respectively. The terms effectively cued

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

both general categorization (retention of the shared element <board>) and subcategorization (addition of different pre-modifying elements). However, (or so the anectodal story goes; cf. http://en.wikipedia.org/wiki/Port_(nautical)) from an acoustic-perceptual perspective, when implemented under harsh climatic conditions at sea, the terms were not distinctive enough - and larboard was gradually replaced by <port>.

7. It is interesting to note that the processes at work are basically the same ones as in fortitions and in general mechanisms of categorization; the difference lies in the application.

8. After Labov's (cf. Labov, 1972) initial study on Martha's Vineyard, he left the topic of how speech features relate to social identities and social values somewhat behind to concentrate on his `attention paid to speech´ model. By drawing attention away from situational contextual factors he aimed at eliciting speaker's vernacular, the assumption being that speech is always monitored to context.

9. When Lambert et al (1960) undertook the task of proving that we primarily evaluate speakers on the basis of their group membership rather than on the individually-based characteristics of their voice (testing in reality the existence of a group-related link between linguistic and social stereotypes), awareness also played a major role. The matched-guise technique was implemented to show that the same (bilingual) speaker was rated quite differently according to the language he or she spoke. The subjects tested thus ignored the fact that they were attributing different sets of group-related psychological attributes to one and the same person in each case and not to different individuals. A panel of judges even rated their own speech variety (French Canadian) as inferior with respect to the more prestigious variety tested (English Canadian) –a result which in all likelihood would not have been obtained if the researchers had implemented direct methods of elicitation.

10. It should be noted that Eckert's reasoning is in line with a variety of models on style shifting opting for approaches which assign a more active role to the speaker. Instead of merely adapting himself lectally to the circumstances of a given situation, a speaker may create personae (Coupland, 2001) or engage in proactive identity construction (Walfram & Schilling-Estes, 1998).

11. It is important to note that the perspective adopted, while still category-internal, does not depict the structure of a phonemic category, but rather language-internal variation. I do not wish to argue, in this respect, that lectal varieties are schematic with respect to the features they are composed of. Just as a category such as BIRD cannot be said to be schematic with respect to features such as eggs, wings, or feathers but rather to members such as robins or penguins, a lectal category such as British English is schematic with respect to Glaswegian or Liverpudlian, but not in a direct way with respect to phonetic features: the features in question point metonymically to a lectal category. Any schematization involved (e.g. from linguistic stereotype to token) is achieved by way of such mechanisms.

REFERENCES

Aitchison, J. (1991). *Language change: Progress or decay?*. 2nd edition. Cambridge: Cambridge University Press.

Bybee, J. (1988). Morphology as lexical organization. In M.l Hammond & M. Noonan (Eds.), *Theoretical mrphology. Approaches to modern linguistics*. San Diego: Academic Press, pp. 119-142.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Bybee, J. (1999). Usage-based phonology. In M. Darnell *et al.* (Eds.), *Functionalism and formalism in linguistics, vol. 1: General papers*. Amsterdam/Philadelphia: John Benjamins, pp. 211-242.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Coupland, N. (2001). Language, situation, and the relational self: Theorizing dialect-style in sociolinguistics. In Penelope Eckert and John. R. Rickford (eds.), *Style and sociolinguistic variation*. Cambridge: Cambridge University Press, pp. 185-210.

Eckert, P. (2004). The meaning of style. Online document: http://studentorgs.utexas.edu/salsa/salsaproceedings/salsa11/SALSA11papers/eckert.pdf

Geeraerts, D. (1997). *Diachronic prototype semantics A contribution to historical lexicology*. Oxford: Clarendon Press.

Geeraerts, D. (2001). On measuring lexical convergence. In A. Soares da Silva (Ed.), *Linguagem e cognição.* Braga: Associação Portuguesa de Linguística, pp. 51-61.

Geeraerts, D. (2005). Lectal variation and empirical data in cognitive linguistics. In F. Ruiz de Mendoza Ibáñez & S. Peña Cervel (Eds.), *Cognitive linguistics: Internal dynamics and inter- disciplinary interaction, Cognitive Linguistics Research 32*. Berlin/New York: Mouton de Gruyter, pp. 163-190.

Geeraerts, D., Grondelaers, S., & Bakema, P. (1994). *The structure of lexical variation: Meaning, naming, and context*. Berlin/New York: Mouton de Gruyter.

Giles, H., Mulac, A., Bradac, J. J., & Johnson, P. (1987). Speech accommodation theory: The next decade and beyond. In M. L. McLaughlin (Ed.), *Communication yearbook 10*. Beverly Hills: Sage, pp. 13-48.

Halliday, M. A. K. (1978). *Language as social semiotic. The social interpretation of language and meaning*. London: Edward Arnold.

Holland, D., & Quinn, N. (1987). *Cultural models in language and thought*. Cambridge: Cambridge University Press.

Irvine, J. T., & Gal, S. (2000). Language ideology and linguistic differentiation. In P. V. Kroskrity (Ed.), *Regimes of language: Ideologies, politics, and identities*. Santa Fe, NM: SAR Press, pp. 35-83.

Jakobson, R. (1960). Closing statement: Linguistics and poetics. In Thomas A. Sebeok (Ed.), *Style in languag*e. Cambridge, Mass.: MIT Press, pp. 350-377.

Kristiansen, G. (2001). Social and linguistic stereotyping: A cognitive approach to accents. *Estudios Ingleses de la Universidad Complutense*, 9, 129-145.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Kristiansen, G. (2003). How to do things with allophones: Linguistic stereotypes as cognitive reference points in social cognition. In: R. Dirven *et al.* (Es.), *Cognitive models in language and thought. Ideology, metaphors and meanings. Cognitive Linguistics Research 24*. Berlin/New York: Mouton de Gruyter, pp. 69-120.

Kristiansen, G. (forthcoming a). Style-shifting and shifting styles. A socio-cognitive approach to lectal variation. In: G. Kristiansen & R. Dirven (Eds.), *Cognitive sociolinguistics. Language variation, cultural models, social systems*. Berlin/New York: Mouton de Gruyter.

Kristiansen, G. (forthcoming b). Idealized cultural models: The group as a variable in the development of cognitive schemata. In: R. M. Frank *et al.* (Eds.), *Body, language, and mind Vol. 2: Cultural situatedness*. Berlin/New York: Mouton de Gruyter.

Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.

Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. 2nd. Ed. New York: Liveright.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Lambert, W. E., Hodgson, R., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, 60, 44-51.

Langacker, R. W. (1988). A usage-based model. In Brygida Rudka-Ostyn (Ed.), *Topics in cognitive linguistics*. Amsterdam/Philadelphia: John Benjamins, pp. 127-161.

Langacker, R. W. (1999). *Grammar and conceptualization*. Berlin/New York: Mouton de Gruyter.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

Mompeán, J. A. (2004). Category overlap and neutralization: The importance of speakers' classifications in phonology. *Cognitive Linguistics*, 15(4), 429-469.

Nathan, G. S. (1986). Phonemes as mental categories. *Proceedings of the Berkeley Linguistic Society*, 12, 212-223.

Nathan, G. S. (1994). How the phoneme inventory gets its shape: Cognitive grammar's view of phonological systems. *Rivista di Linguistica*, 6, 275-287.

Nathan, G. S. (1996). Steps towards a cognitive phonology. In B. Hurch & R. Rhodes (Eds.), *Natural phonology: The state of the art.* Berlin/New York: Mouton de Gruyter, pp. 107-120.

Nathan, G. S. (1999). What functionalists can learn from formalists in phonology. In M. Darnell *et al.* (Eds.), *Functionalism and formalism in linguistics, vol. 1: General papers*. Amsterdam/Philadelphia: John Benjamins, pp. 305-327.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Niedzielski, N., & Preston, D. (2000). *Folk linguistics*. Berlin/New York: Mouton de Gruyter.

Nunberg, G. (1978). *The Pragmatics of reference*. Bloomington, Indiana: Indiana University Linguistics Club.

Nuyts, J. (2005). Brothers in arms? On the relations between Cognitive and Functional Linguistics. In F. J. Ruiz de Mendoza Ibáñez & M. S. Peña Cervel (Eds.), *Cognitive linguistics. Internal dynamics and interdisciplinary interaction, Cognitive Linguistics Research 32*. Berlin/New York: Mouton de Gruyter, pp. 69-100.

Orton, H., Sanderson, S., & Widdowson, J. (Eds.) (1978). *The linguistic atlas of England*. London: Croom Helm.

Purnell, T., Idsardi, W. J., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, 18, 10-30.

Sandra, D., & Rice, S. (1995). Network analyses of prepositional meaning: Mirroring whose mind – the linguist's or the language user's?. *Cognitive Linguistics*, 6(1), 89-130.

Sharifian, F. (2003). On cultural conceptualisations. *Journal of Cognition and Culture.* 3(3), 187-206.

Stampe, D. (1979). *A dissertation on natural phonology*. New York: Garland Publishing.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations*. Monterey, CA: Brooks/Cole, pp. 33-47.

Taylor, J. R. (1990). Schemas, prototypes, and models: In search of the unity of the sign. In S. L. Tsohatzidis (ed.), *Meanings and prototypes: Studies in linguistic categorization*. London: Routledge, pp. 521-534.

Taylor, J. R. (1995). *Linguistic categorization. Prototypes in linguistic theory*. 2nd. edition. Oxford: Clarendon Press.

Taylor, J. R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.

Tuggy, D. (1993). Ambiguity, polysemy, and vagueness. *Cognitive Linguistics*, 4(3), 273-290.

van Bezooijen, R., & Gooskens, Ch. (1999). Identification of language varieties. The contribution of different linguistic levels. *Journal of Language and Social Psychology*, 18(1), 31-48.

Werner, H. (1934). L'unité des sens. *Journal de Psychologie Normale et Pathologique*, 31, 190-205.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Werner, H. (1957). *Comparative psychology of mental development*. New York: International Universities Press.

Werner, H., & Wapner, S. (1952). Toward a general theory of perception. *Psychological Review*, 59: 324-38.

Wolfram, W., & Schilling-Estes, N. (1998) *American English: Dialects and variation*. Oxford: Blackwell.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

The Phoneme as a Basic-Level Category: Experimental Evidence from English[1]

JOSÉ A. MOMPEÁN
University of Murcia[*]

ABSTRACT

This paper presents the results of a *concept formation* experiment that provides evidence on the possible existence of a basic-level of taxonomic organization in phonological categories as conceived of by phonetically naïve, native speakers of English. This level is roughly equivalent to the phoneme as described by phonologists and linguists. The reason why the phoneme could be considered as the basic level of taxonomies of phonological categories is discussed.

KEYWORDS: classical view of categorization, taxonomies, basic level, phonological categories, concept formation.

[*] ***Address for correspondence***: José A. Mompeán. Departamento de Filología Inglesa. Facultad de Letras. Campus de *La Merced*. Universidad de Murcia, 30071. Murcia, Spain. Tel. 00 34 968 364383, Fax: 00 34 968363185; e-mail: mompean@um.es

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I. HIERARCHICAL ORGANISATION OF CONCEPTUAL CATEGORIES

One basic characteristic of human conceptual categories and concepts is that they do not exist independently of one another in memory. Instead, they tend to be organised into systems where they are related to one another in various ways. In this respect, it has been claimed that most (if not all) of our cognitive categories are hierarchically organised (Neisser & Weene, 1962: 640) and the most typical type of hierarchical organisation is, in turn, the taxonomic one. In taxonomies, categories are organised by the 'type' relation, which specifies that one category is a type or kind of another. Thus, the category WHITE-TAILED SEA EAGLE (which is instantiated by many different white-tailed sea eagles in the real world) is a type of SEA EAGLE, which is in turn a type of EAGLE and this a type of BIRD. The category BIRD is a type of ANIMAL and this a type of LIFE FORM.

Hierarchical taxonomic organisation has been the focus of a great deal of experimental research in the cognitive sciences for the past fifty years. In this body of research, taxonomic organisation is usually conceived of as a vertical axis in which there are different levels of abstraction, occupied by different contrasting categories. These levels of abstraction are more inclusive as we move upwards and less inclusive as we move downwards. Thus each category within a taxonomy includes others (unless it is the lowest level category) or is included in others (unless it is the highest level category). For example, the categories WHITE-TAILED SEA EAGLE, BALD SEA EAGLE and WHITE-BELLIED SEA EAGLE are included in the category SEA EAGLE, which is in turn included, together with other categories like GOLDEN EAGLE or HARPY EAGLE in the category EAGLE. The category EAGLE is included, alongside categories like ROBIN, FLAMINGO, etc. in the category BIRD and the latter, with categories like REPTILE, MAMMAL or INSECT in the category ANIMAL. The category ANIMAL, with others like PLANT or BACTERIA are included in the higher-level category LIFE FORM.

One important fact about empirical research on hierarchical taxonomic organisation is that for that it has for centuries been conceived of according to the so-called 'classical' Aristotelian theory of categories and categorisation, which claims that categories are discrete entities characterized by a set of properties shared by all their category members and necessary and sufficient to establish category membership. According to the classical view, categories should be clearly defined and mutually exclusive (any given entity of a given classification universe belongs unequivocally to one, and only one, of the proposed categories). This view has typically been the dominant position in philosophy (see e.g. Margolis, 1994; Rey, 1983, 1985; Sutcliffe, 1993 for recent defences), psychology, with the

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

traditional concept formation and learning experiments of the behaviourists initially (e.g. Hull, 1920) and the information processors later (e.g. Bourne, 1966, 1970; Bruner *et al.*, 1956; Hunt, 1962), linguistics (e.g. Chomsky & Halle, 1968; Katz & Postal, 1964), etc. The classical view of categorization provides an intuitively appealing account of classification and the nature of conceptual structure. However, the view has been running into numerous problems since experimental research started to be conducted. These problems include the classical view's failure to account for the lack of defining features for many categories, the use of non-necessary features in categorisation by subjects, the existence of unclear category members, or the phenomenon of typicality and typicality effects (see Mompeán, 2002 for a review of these problems).

Another problem of the classical view of categories and categorisation that empirical research has revealed is the inconsistency of many taxonomic phenomena with the classical theory's view. For instance, classical taxonomies (particularly scientific system) appear to have an excessive number of levels from the non-scientist, everyday person's point of view (Ungerer & Schmid, 1996: 64). In modern Linnaean taxonomies, for instance, all species can be simply classified in a ranked taxonomy starting with *domains*, and the latter into *kingdoms*. Kingdoms are divided into *phyla* (for animals) or *divisions* (for plants). Phyla and divisions are divided into *classes*, and they, in turn, into *orders*, *families*, *genera*, and *species*. However, between these levels there are many others, which have been added in certain disciplines whose subject matter is replete with species requiring classification (see URL 1 for a review). In contrast, people's folk taxonomies are often not so elaborated. Thus, in anthropological and ethnobiological taxonomic studies (e.g. Berlin, 1978, 1992; Berlin *et al.*, 1973; Brown *et al.*, 1976), five levels are often posited: *unique beginner*, *life form*, *generic*, *specific*, and *varietal*. In cognitive psychology (e.g. Rosch, 1978; Rosch *et al.*, 1976) and related disciplines, three levels are often set up: *superordinate* (which includes *unique beginner* and *life form*), *basic*, and *subordinate* (which includes *specific* and *varietal*). However, most authors implicitly or explicitly acknowledge that any taxonomy is best considered as a continuum of differentiation with no strict layers on which any category falls (Murphy & Brownell, 1985; Rosch *et al.*, 1976; Tanaka & Taylor, 1991; Ungerer & Schmid, 1996).

A further problem of the classical theory's view on taxonomic organisation is the well-known finding that there seems to be, in conceptual taxonomies, a particular level of specificity that enjoys psychological salience or primacy. This is the *generic* level (in ethnobiological terms) or *basic* (in cognitive psychological terms) level. For example, the basic level of abstraction in the hierarchy that goes (from top to bottom), from LIFE FORM to

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

WHITE-TAILED SEA EAGLE, is the category BIRD. For a large number of years, investigators have shown that the basic level has a special status in a variety of tasks,[2] which makes no sense in the classical theory's view on hierarchical structure, in which no particular level of abstraction should have a special status (Mervis, 1980: 285).

The special centrality of the basic level is revealed in tasks that reflect the contents of category knowledge, inference drawing and recall/recognition memory. In this respect, the basic level (e.g. CHAIR, CAR, DOG) is the level at which subjects list more attributes for category members (e.g. Horton & Markman, 1980; Mervis & Greco, 1984; Mervis & Rosch, 1981; Murphy & Brownell, 1985; Rosch *et al.*, 1976).[3] The richer attribute structure or feature information that basic-level categories possess may be the reason why the basic level is also the level at which more inferences can be drawn -particularly in comparison with superordinate categories- (Gelman & Markman, 1986; Gelman & O'Reilly, 1988). Finally, the basic level is the preferred level for retaining episodic information in memory that is used later for recall. Thus, subjects presented with either subordinate terms (e.g. sports car) or superordinate terms (e.g. vehicle) tend to falsely report basic-level terms (e.g., car) instead (Pansky & Koriat, 2004).

The special salience of the basic level is also revealed in tasks that involve people's perception of objects or the mental capacity to image them. In this respect, the basic level is the highest level in which category members have similar overall perceived shapes so that, as a consequence, the average shape of a number of instances of basic-level categories like CHAIR, for instance, are still recognisable or identifiable as an instance of that category (e.g. Rosch, 1978; Rosch *et al.*, 1976).[4] The basic level is also the highest level at which it is possible to form a relatively concrete mental image of an average member of the category (in the absence of that object) which is isomorphic to an average category member, an ability known as "imaging capacity" (Bolton, 1977: 56). As a case in point, people have mental images of basic-level categories like CHAIR but they do not have abstract mental images of superordinate categories like FURNITURE that are not images of basic-level objects like chairs, tables, beds, etc. (Rosch, 1978; Rosch *et al.*, 1976).

Further tasks reflect people's motor or verbal behaviour towards members of categories. In this respect, the basic level is the highest level in a taxonomy at which a person uses similar motor actions for interacting with category members (Rosch *et al.*, 1978).[5] In addition, basic-level categories (and basic-level category names) are primarily used when identifying objects in controlled context-free free-naming tasks (e.g. Jolicoeur *et al.*, 1984; Murphy & Brownell, 1985; Murphy & Wisniewski, 1989a; Rosch *et al.*, 1976; Smith *et al*, 1978; Tanaka

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

& Taylor, 1991).[6] In addition, such identifications are usually faster at the basic level than at any other abstraction level. In more naturalistic situations like normal everyday conversation, basic-level category names are also more frequently used (Anglin, 1977; Berlin *et al.* 1973; Brown, 1958, 1976; Cruse, 1977; Downing, 1977).[7]

Finally, it seems that basic-level categories are, throughout develoment, often learned first (basic-level names certainly are) and that they are easier to learn in experimental situations (see e.g. Callanan, 1985; Horton & Markman, 1980; Mervis, 1987; Mervis & Crisafi, 1982; Rosch *et al.*, 1976; Waxman *et al.* 1991).[8]

II. HIERARCHICAL ORGANISATION OF PHONOLOGICAL CATEGORIES

In view of the extensive amount of research carried out on different sorts of conceptual categories in general (mainly of a visual or semantic type) and on taxonomic organisation of such categories in particular, the approach can be explored of whether phonological categories are also hierarchically structured. This approach rests upon two assumptions. The first assumption is that people actually group sounds into categories or, as Nathan claims (1996: 112), that "…sounds... are categorized in the same way as all other things in the world are". This, however, is a well-established fact as shown by the long history of research in the fields of speech perception and experimental phonology where, with different techniques like phoneme monitoring (see e.g. Foss, 1998), absolute identification and differential discrimination -in either its same-different or ABX versions- or concept formation (see e.g. Weitzman, 1993) have shown that speakers can group sounds into categories and use these categories for further processing and interaction. The second assumption is that people's ability to categorise sounds into categories may result in the creation or formation of conceptual categories to which technical concepts used by linguists like 'phoneme', 'fricative', etc. more or less correspond.

The two assumptions mentioned above have traditionally been uncommon in the history of linguistics but they are central in cognitive linguistics (Fraser, 2006), where the view has long been held (mainly in relation to phonemic categories) that phonological terms also refer to conceptual categories (or concepts) in the sense that speakers can assign phonetically different sounds to them and draw inferences based on them (e.g. Fraser, 2006; Mompeán, 2004; Nathan, 1986, 1996; Taylor, 2002, 2003, 2006). Therefore, if language users can form metalinguistic phonological categories, the latter can probably be related to one another

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

hierarchically. Again, this approach has only been explored in cognitive linguistics. Thus, Taylor (2002: 145-150) discusses a plausible taxonomy of phonological segments with the superordinate category SEGMENT at the top of the hierarchy and lower levels of abstraction represented by categories like VOWEL and CONSONANT (at a little lower level than the category SEGMENT), PHONEME categories (further down) and ALLOPHONE categories at the bottom of the taxonomy. Taylor (2002: 149-150; 2006: 44), Nathan (2007) and Mompeán (1999) go further to suggest that, in a taxonomy of phonological concepts, the level of the traditional phoneme probably has some basic-level status. However, empirical studies on the issue are almost inexistent. The only exception (to the author's knowledge) is Jeri J. Jaeger's doctoral dissertation (Jaeger, 1980, also summarised in Jaeger, 1986). Using the experimental paradigm known as *concept formation*, Jaeger found that the percentage of adult subjects who formed the category PHONEME K (i.e. /k/) was higher –100% in her CF experiment -and the number of trials to criterion was fewer- than the number of subjects who formed 'feature' categories like -ANTERIOR, +ANTERIOR, +SONORANT, and +VOICE, exemplified each by different phonemes and learned, respectively, by 79%, 50%, 50% and 38%-50% of experimental subjects (the number of trials was also higher in the 'feature' categories). Jaeger interpreted her results as evidence that phonemes were basic-level categories as compared to feature categories.

Pioneering and insightful as Jaeger's work is for the study of taxonomic hierarchies in phonology, her work seems to have failed to fully outline the structure of a taxonomic hierarchy for the categories she studied. For instance, if feature categories of the type that she studied are subordinate categories (taking for granted that phoneme-sized categories are the basic-level), then by the 'type' relation taxonomies are based on, all the member of a given 'feature' category should also be members of a single higher-level (basic-level) phoneme category. However, this is not true for the categories she studied. Thus, for the category +VOICE, Jaeger included, as positive tokens, words which began with [v, ð, z, m, n, r, w, j, l]. Clearly, not all the sounds of that category instantiate a single phoneme category. This means that 'feature' categories (which may be less salient than phoneme categories as Jaeger showed) should be best conceived of as superordinate, not subordinate categories. Subordinate categories can be allophones, but Jaeger did not look at these categories nor to categories higher in a taxonomy such as CONSONANT or SPEECH SEGMENT. In addition, Jaeger included in her discussion references to an experiment where the category learned was Vowel Shift alternations of the type *serene-serenity*, *divine-divinity*, etc. popularised by Chomsky &

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Halle (1968), for which speakers scored relatively high (73% of speakers formed the category) and she later claimed that "….phoneme-based categories appear to be at a more basic level of abstraction for English speaking subjects than do the phonetic features-based categories, with the rule-based category somewhat intermediate" (Jaeger, 1980: 372) and that "…the phonemic level is the basic level of categorisation for speech sounds, and features are a subordinate level" (p. 381). The Vowel Shift rule category is clearly to be excluded from a taxonomy of speech sounds that includes phonemes, allophones or features at different levels of abstraction. An allophone can be an instance of a phoneme and the latter an instance of a 'feature' category, but neither of these, by themselves, can be an instance of a category that is relational and includes two phonemes.

Another problem of Jaeger's discussion of taxonomic organisation is her confusion between taxonomic and partonomic hierarchies. Based on her experimental research, Jaeger claims that, for English speakers, the phoneme has basic-level status in taxonomic phonological hierarchies whereas "…. syllables, words, etc. are superordinate levels", which is not the case of Japanese speakers, for whom "…the syllable is the basic level of categorization of the sounds of their language; words, etc. are superordinate, and phonemes, features, etc. are subordinate" (Jaeger, 1980: 146). However, units like the syllable, the phonological word, etc. should not be brought up in discussions of the taxonomic organization of categories since such terms refer to a conceptual organization of a hierarchical, but not taxonomic type. As Taylor warns (2002: 149), "...the relation between the syllable and phoneme is not a schema-instance relation, but the relation of a whole to a part". In fact, there has also been research on partonomic organisation and partonomies (see e.g. Tversky, 1989, 1990; Tversky & Hemenway, 1984, 1991) which are organised by the 'part relation', which specifies that one concept represents a part of another. Thus, in the well-known *body part* partonomy, a finger is a part of a hand, which is a part of the arm, which is a part of the body. Fingers are not included in the class of hands, which are not included in the class of bodies. Similarly, the kind of relationship that exists between phonemes, rhymes, syllables, and phonological words (this list is not exhaustive), is of a partonomic nature (a phoneme is a part of an onset or rhyme, which are a part of a syllable, which is a part of a phonological word, etc.).[9]

Given the small empirical evidence available on phonological taxonomic organisation, the main aim of this paper is to provide some empirical evidence on speakers' ability to categorise sounds at different levels of abstraction that can be taxonomically related and find out whether any of the levels of abstraction has greater salience (or basic-level status). More

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

specifically, the research questions investigated in this paper are: a) can speakers categorise sounds at different levels of abstraction that can be taxonomically related? and b) if so, is there any level of abstraction that is more salient than the others?

Based on all the research discussed above, the hypotheses entertained in this paper are that: a) speakers will be able to categorise sounds at different levels of abstraction that can be taxonomically related; and b) that some evidence of a basic level of abstraction in phonological taxonomies will also emerge. To test these hypotheses, four different experiments were conducted in order to investigate four categories that, from now on, will be referred to as CONSONANT, PLOSIVE, PHONEME P, and ASPIRATED P. This study is based on the assumption that these categories are taxonomically organised, which further assumes that any given sound can be cross-classified (a given sound can be an instance, for instance, of the categories ASPIRATED P, PHONEME P, PLOSIVE, and CONSONANT at the same time).

III. METHOD

III.1. Participants

Eighty native speakers of English between the ages of 18 and 45 (mean age 23 years) took part in the experiments reported below. There were 38 men and 42 women. The subjects were recruited on the University of Murcia campuses or in the town of Murcia through advertisements. None of them had received formal instruction in phonetics and/or phonology in the past and all of them had reached university. For this reason, the whole group could be described as educated (and so fully literate in English) but phonetically naïve. Subjects reported no history of a speech or hearing disorder.

III.2. Apparatus

All the experimental events in the experiments reported below were controlled by a computer in which a software implementation of the experimental technique called *concept formation* (henceforth CF) had been installed.[10]

The CF technique consists of a *training* session followed by a *test* session (see Jaeger, 1986; Mompeán, 2002, for a full overview of the specifics of the technique). The aim of the *training* session is to 'teach' the experimental subjects a phenomenon under investigation. This is done by training them to classify a (usually large) set of items into different groups or

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

categories that have been pre-defined by the experimenter. Thus, subjects are trained to respond to a particular type of stimuli that exemplifies a given category (i.e. positive stimuli) in one way, and to respond to another type of stimuli that does not exemplify that category (i.e. negative stimuli) in a different way. In the training session there are three critical events: stimulus presentation, response, and informative feedback. These three events, occurring in that order, constitute one trial on the problem.[11] After each stimulus is presented, and the subject has some notion of what the category involves, the subject's task consists in trying to give the correct response (as instructed) after which the actual correct response is indicated with the provision of feedback. Feedback informs subjects about the status of each instance they are exposed to (whether it is a positive token of the to-be-formed category or not).

In the *test* session, the subject's task is the same as in the previous one except there is no feedback because an aim of this session is to find out whether the subject has actually guessed what the target category was. In principle, if the subject reached criterion in the training session, s/he should have no problems in continuing to provide correct responses to positive and negative stimuli of the type presented in the training session. However, to guarantee that the subject has actually found out what category the experimenter had in mind, the test session also makes use of the so-called *control* tokens (positive or negative instances of the category not yet encountered by the subject), which are checks on the possibility that the subject has not formed a category different from that intended by the experimenter, or that s/he may have just memorized the members of the category encountered in the training session. If the subject generalizes his/her responses to these new cases correctly, the classificatory behaviour more clearly indicates that the subject has actually 'formed' the category.

III.3. Stimuli

The stimuli used in the present study consisted of 400 monosyllabic English words (100 per experiment), produced by a 22-year-old female native RP speaker of English from the south of England.

In the training sessions of each experiment, there were 32 positive and 28 negative items. The negative tokens also included *interfering* and *non-interfering* items. Interfering items in this study were those containing potential orthographic and/or phonetic interference. In the test sessions there were 19 positive, 12 negative (some of them controls) and 9 test

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

tokens (not analysed for the present study),[12] in all experiments except for experiment 4, where no test items were included and 22 positive and 18 negative tokens were used instead.

The positive stimuli of both the training and the test sessions exemplified, for the category CONSONANT, word-initial instances of plosives, fricatives, affricates and nasals while negative stimuli were words beginning with a vowel. For the category PLOSIVE positive stimuli were words beginning with oral plosives while negative stimuli consisted of words beginning with fricatives and nasals. The category PHONEME P was instantiated by different allophonic realisations of /p/ in pre-nuclear and post-nuclear positions with different types of release, degree of aspiration, etc. while negative stimuli did not contain any realisation of /p/. The category ASPIRATED P was exemplified by aspirated pre-nuclear realisations of /p/ (before a vowel or a devoiced approximant) while negative items included pre-nuclear and post-nuclear realisations with inaudible release, masked release, weak (if any) aspiration, etc.

As far as (negative) interfering items are concerned, orthographic interference was considered likely in words containing letters that are typical spellings of the target sounds but that are silent or have phonetic values other than their prototypical ones in those words. For instance, in experiment 2 (category PLOSIVE) a word beginning with <ps> like *psalm* was considered potentially interfering since the letter <p>, a typical representation of the voiceless bilabial plosive, is a silent letter. Phonetic influence may derive from the presence of phonetically similar sounds to the ones that instantiate the target category but that are not to be included in the category. For instance, in experiment 3 (category PHONEME P), it was considered that phonetic interference could be caused by the presence of /b/ in word-initial position as it is partially or wholly devoiced in that position. As far as controls are concerned, these included, for instance, phonemes not previously found in the training session (e.g. /t/ in the category PLOSIVE) or allophonic realisations not previously encountered (e.g. in the category ASPIRATED P).

The Appendix at the end of this paper contains the actual list of words used and their category status (positive, negative, interfering, non-interfering, control, test) for each of the four experiments carried out.

III.4. Procedure

All the CF experiments were conducted in a sound-attenuated room and experimental events were controlled by a computer in which a software program specifically designed to perform

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

the CF experiments had been installed (see Mompeán, 2002 for details). The experimental events were monitored by an experimenter on-line from an adjacent room.

For this study, the 80 subjects were randomly assigned to one of four groups (20 per experiment). Subjects were given a sheet of instructions asking them to perform a task in which they had to focus on the sounds of words, never the spelling. The instructions told them that some words had "...a certain type of sound in the initial position of the word..." (exp. 1 & 2), "...a certain type of consonantal sound somewhere in the word…" (exp. 3), or that all words contained basically the 'same' consonantal sound but that some examples of the consonantal sound had "certain characteristics" (exp. 4).

The instructions also told the subjects that after listening (over headphones) to each word (only once), they would be provided with an answer as to whether or not the word had included the to-be-identified type of (consonantal) sound. Red/green rectangles on the screen of the computer would then disappear/remain on the screen depending on whether the words presented contained/lacked the to-be-identified type of sound. The instructions also told the subjects to begin responding as soon as they heard each new word once they had some idea of what the target type of sound was. The response was carried out by pressing either a red or a green key on the keyboard. Subjects were also informed that after a certain number of trials, feedback would be no longer provided (though they would be told when feedback provision would stop).

The training session began only when the experimenter was sure, through a short conversation after the subjects read the instructions, that the subject had understood the instructions well. Subjects were run individually.

IV. RESULTS

The results show that the four categories investigated were positively 'formed' by over 50% of the experimental subjects in each experiment, ranging from 60% of speakers (category CONSONANT) to 100% (categories PHONEME P and ASPIRATED P). The results also show that not all categories were equally salient or as easy to form. The measures gathered in this study giving evidence about the difficulty of the categories are: 1) the number of subjects who reached criterion in the training session, 2) the average number of correct responses in the training session, 3) the standard deviation (and range) of the individual scores of correct responses in the training session, and 4) the percentages of correct responses to positive and negative stimuli in both the learning and test sessions of each experiment. The order in which

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

positive, negative and test tokens were presented in the four experiments was constant so the results are comparable across the four experiments.

The first three pieces of evidence are included in Table 1. Subjects were considered to have 'formed' a given category if they reached 37 correct responses, which guaranteed that their classifying behaviour had not been random in the training session (P-value 0.03 < 0.05). Given this, the number of criterial subjects in the four CF experiments was 12 (exp. 1), 14 (exp. 2) and 20 (exp. 3 & 4). The table also shows that the average number of correct responses in the training session was 48.67 (range: 37-59, s.d. = 6.50) in exp. 1, 47 (range: 37-59, s.d. = 7.06) in exp. 2, 56.55 (range 51-60, s.d.= 2.06) in exp. 3, and 51.4 (range 37-59, s.d. = 5.15) in exp. 4.

Category (experiment)	**Criterial subjects**	**Mean correct responses (Training session)**	**Range & Standard Deviation**
CONSONANT (exp. 1)	12	48.67	Range: 37-59. s.d= 6.50
PLOSIVE (exp. 2)	14	47	Range: 37-59. s.d= 7.06
PHONEME P (exp. 3)	20	56.55	Range: 51-60. s.d= 2.06
ASPIRATED P (exp. 4)	20	51.4	Range: 37-59. s.d= 5.15

Table 1: Criterial subjects per category and experiment and category, subjects' mean correct responses (maximum 60) in the training session, range and standard deviation.

The number of correct, incorrect, and null responses to positive, negative and total stimuli as well as percentages of correct responses to each stimulus type in the training session are shown in Table 2.

Category	**Stimulus type**	**Type of response**			**% correct responses**	**Items**	**Responses elicited**
		C	**I**	**N**			
CONSONANT	Positive	322	33	29	83.85%	32	284
	Negative	286	30	20	85.12%	28	336
	Total	608	63	49	84.44%	60	720
PLOSIVE	Positive	364	74	10	81.25%	32	448
	Negative	290	98	4	73.98%	28	392
	Total	654	172	14	77.86%	60	840
PHONEME P	Positive	598	12	30	93.44%	32	640
	Negative	532	12	16	95%	28	560
	Total	1130	24	46	94.17%	60	1200
ASPIRATED P	Positive	554	27	59	86.56%	32	640
	Negative	473	39	48	84.46%	28	560
	Total	1027	66	107	85.58%	60	1200

Table 2: Category studied, number of correct (C), incorrect (I) and null (N) responses and percent correct responses to positive, negative and total stimuli, items and responses elicited (training session)

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

This table also shows the number of items per type of stimulus, and the number of responses elicited (which results from multiplying the number of items by the number of subjects who reached the established criterion in the training session). The equivalent information obtained from subjects' performance in the test session is shown in Table 3.

As tables 2 and 3 show, in all four experiments, the percentages of correct responses to both positive and negative stimuli substantially increase in the test session as compared with the training session. In other words, correct responses (see also Table 4) were significantly more frequent in the test session than in the training session as shown by respective contrasts of proportions (exp. 1 -CONSONANT-: 84.44% vs. 93.01%; exp. 2 -PLOSIVE-: 77.86% vs. 89.40%; exp. 3 -PHONEME P-: 94.17% vs. 98.71%; exp. 4 -ASPIRATED P-: 85.58% vs. 97.5%; p-value: $0.000 < 0.05$).

Category	Stimulus type	Type of response			% correct responses	Items	Responses elicited
		C	I	N			
CONSONANT	Positive	218	2	8	95.61%	19	228
	Negative	128	15	1	88.89%	12	144
	Total	346	17	9	93.01%	31	372
PLOSIVE	Positive	248	13	5	93.23%	19	266
	Negative	140	25	3	83.33%	12	168
	Total	388	38	8	89.40%	31	434
PHONEME P	Positive	379	0	1	99.74%	19	380
	Negative	233	2	5	97.08%	12	240
	Total	612	2	6	98.71%	31	620
ASPIRATED P	Positive	434	5	1	98.64%	22	440
	Negative	346	10	4	96.11%	18	360
	Total	780	15	5	97.5%	40	800

Table 3: Category studied, number of correct (C), incorrect (I) and null (N) responses and percent correct responses to positive, negative and total stimuli, items and responses elicited (test session)

Category (experiment)	Stimulus type	% correct responses	
		Training session	Test session
CONSONANT	Positive	83.85%	95.61%
	Negative	85.12%	88.89%
	Total	84.44%	93.01%
PLOSIVE	Positive	81.25%	93.23%
	Negative	73.98%	83.33%
	Total	77.86%	89.40%
PHONEME P	Positive	93.44%	99.74%
	Negative	95%	97.08%
	Total	94.17%	98.71%
ASPIRATED P	Positive	86.56%	98.64%
	Negative	84.46%	96.11%
	Total	85.58%	97.5%

Table 4. Percentages of correct responses to positive, negative, and total stimuli (training and test sessions).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

The results clearly show that the category PHONEME P (exp. 3) was the easiest to form of the four categories studied. PHONEME P and ASPIRATED P were the only categories for which all experimental subjects reached criterion. However, subjects in experiment 3 made more correct responses in the training session as an average than subjects in any of the other three experiments, the scores of the different subjects (i.e. range 51-60) were higher and differed less than those of the experimental subjects in the other experiments and the standard deviation of those scores (i.e. 2.06) was also the lowest. In addition, the subjects in experiment 3 made more correct responses to both positive and negative items in both the learning and the test sessions than the subjects in experiment 4. Thus, although the 20 experimental subjects of experiments 3 and 4 formed the categories PHONEME P and ASPIRATED P respectively, subjects in experiment 3 performed better than those in experiment 4 and much better than in experiments 1 (category CONSONANT) and 2 (category PLOSIVE).

V. DISCUSSION

In retrospect, it is not at all surprising that the categories CONSONANT and PLOSIVE, based on criteria like degree of constriction of the oral tract (plus velic closure/opening in the case of PLOSIVE), were more difficult to form than the categories ASPIRATED P and PHONEME P. Previous studies have found that 'feature' categories (instantiated by different speech segments that do not belong to the same segment-sized category) are more difficult to form than categories instantiated by speech segments that are classified as members of the same phoneme category (according to adult standards). Jeri J. Jaeger's (1980) CF experiments discussed above are a example of this. John Ohala's (1986) study is also revealing. In this latter work, Ohala taught one group of adult English-speaking subjects to group the [k] in a word like *school* with [kʰ] as in *cool*, and he taught another set of subjects to group [k] with [g], as in *ago* and [g̊], as in *good*. The first experimental group formed the category easily but many subjects in the second group could not form it at all, and those who did described the category in a disjunctive way (e.g. "[gə] sounds or the [kʰə] sound after *s*"). According to Ohala, the findings revealed that [kʰ, k] was more likely to be a pre-established grouping for subjects than [g̊, g, k], which is based on features like 'velar', 'stop' and 'oral', but whose instances do not belong to a single phoneme category for the experimental subjects as in the case of [kʰ, k]. As another case in point, Fodor and his co-workers (Fodor *et al.*, 1975) found that infants grouped syllables beginning with /p/ (e.g. /pi/, /pu/) more readily than syllables

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

sharing phonetic features like 'voiceless', 'plosive', and 'oral', as is the case of /pi/ and /ka/, but not grouped in a segmental phoneme-sized category according to adult standards. Finally, it should be mentioned that the fact that 'feature'-based categories are more difficult to form than segment-based categories also explains why allophonic categories like ASPIRATED P and PHONEME P were easier to form than the categories CONSONANT and PLOSIVE.

The greater salience of PHONEME P in the present study over the categories PLOSIVE and CONSONANT and the relatively greater salience over the category ASPIRATED P seems to suggest that the category may have some sort of basic-level status in taxonomies of phonological categories for phonetically naïve subjects. This suggestion is based on the fact that in learning tasks in general and CF experiments in particular, basic-level categories are easier to form than non-basic-level categories (Jaeger, 1980: 366), which has already been shown for different sorts of categories other than phonological ones in cognitive and developmental psychology (see references in Section I). If this is so, and given the 'type' relation on which taxonomies are based, allophones would be subordinate categories and 'feature' categories would be superordinate categories (see e.g. Figure 1).

Level	**Category**
Superordinate	CONSONANT PLOSIVE
Basic	PHONEME P
Subordinate	ASPIRATED P

Figure 1: Plausible taxonomic organisation of the categories CONSONANT, PLOSIVE, PHONEME P, and ASPIRATED P

If the phoneme level is actually the most salient level of abstraction in taxonomies of phonological categories for subjects literate in an alphabetic writing system (like the experimental subjects that took part in this study), an explanation of why this could be so is called for. In this respect, the literature on taxonomic organisation mentions two main types of determinants of basic-level status. On the one hand, the basic level is often determined by the structure of the world as it is perceived and processed by cognitive systems (see e.g. Corter & Gluck, 1992; Jolicoeur *et al.*, 1984; Jones, 1983; Lin *et al.*, 1997; Mervis & Rosch, 1981; Rosch, 1978). On the other hand, the basic level also depends on general cultural significance (Berlin, 1992; Berlin *et al.*, 1973; Dougherty, 1978; Stross, 1973) as well as on individual

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

familiarity, expertise or knowledge (Honeck *et al.*, 1987; Medin *et al.*, 1997; Tanaka & Taylor, 1991). In short, the basic level is determined both by the structure of the world and by the contributions of the human perceiver or categoriser like his/her goals, culture, expertise, knowledge, etc. and both types of factors typically interplay to define, for a given subject or population of subjects, the basic level in a given taxonomy (Dougherty, 1978; Mervis, 1980; Rosch *et al.*, 1976).

What 'structural' factors could make the phoneme level have basic-level status? Following a well-known structural explanation of basic-level status that claims that the basic level achieves the optimal balance between informativeness and distinctiveness, the basic level is the level at which categories maximize within-category similarity (i.e. relatively many properties are shared by all category members) while minimizing between-category similarity (i.e. relatively few properties are shared by non-members), attaining optimal cognitive economy (Mervis & Rosch, 1981; Rosch et al., 1976) or cognitive efficiency (Murphy, 1991a, b). Given this, the phonemic level could have basic-level structure because the members of the category (i.e. allophone categories) tend to be more structurally (phonetically) similar to one another than members of higher order 'feature' categories like PLOSIVE, which are based on a single feature or a few features but whose members differ significantly in other important feature specifications (e.g. voicing, place of articulation, aspiration, etc.).[13] This kind of structural similarity tends to make phoneme categories more stable, maximising informativeness (Taylor, 2002: 150): words can be distinguished from one another simply by a change in the phonemic specification of the word. In this respect speakers, even illiterates, are very good at minimal pairs discrimination (see e.g. Adrian *et al.*, 1995; Loureiro *et al.* 2004). It is also the case that the structural similarity of the members of allophonic categories may be as high as that in phoneme categories. However, the gain in informativeness is at the cost of a loss in distinctiveness, which is why listeners are not generally aware of allophonic variation (Abercrombie, 1967: 85, 87; Kreidler, 1989: 98; O'Connor, 1973: 121) and can only be so with special phonetic training (Donegan & Stampe, 1979: 162-164; Nathan, 1996: 112, 1999: 312-313; Pike, 1943: 115; etc.) or why it can be hypothesised that allophone categories will not develop conceptually unless they are somehow perceptually salient, like flaps or glottal stops in English.

Moreover, phonemes are the highest level at which speakers can kinaesthetically sense (in the absence of any audible production) the articulatory movements of an average category member, which seems to relate to Taylor's (2002) assertion that phonemes are the highest units for which speakers can conceptualise or "…. bring to mind an image of the /p/-phoneme

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

in terms of its sound and its articulatory parameters…" (p. 150) and is compatible with the finding that the basic level is the highest level in a taxonomy at which a person uses similar motor actions or movements for interacting with category members (Rosch *et al.*, 1978). In contrast, fewer similar motor actions are used to interact with members of superordinate categories so speakers "… can hardly conceptualise a schematic stop, even less a schematic obstruent…" (p. 150) and, although speakers behave in a very similar way with members of subordinate categories, no more movements are made in common to subordinate than to basic level categories (Rosch, 1978; Rosch *et al.*, 1976).

What cultural or knowledge factors could make the phoneme level have basic-level superiority? The alphabetic writing system seems to be to a great extent responsible for the higher conceptual salience of phonemes. As Taylor (2002: 149) or Nathan (2007) point out, alphabetic writing systems are based on the salience of this level and they never represent sub-phonemic variants by distinct symbols. Thus, the writing system of a language like English represents the phonemic level (although more imperfectly than other languages like Spanish or Turkish) but does not reflect allophonic variation or higher-order phonetic and/or phonological relations, the conceptual salience of the phonemic level is increased. When mastering an alphabetic writing system, speakers "…must realize that *cat*, *act*, and *tack* contain the 'same sounds' arranged in different sequences" (Taylor, 2002: 149), equivalence across phonetic contexts (Pierrehumbert, 2003: 118) being the key characteristic of phonemes (Taylor, 2006). Once subjects have mastered alphabetic writing and the principle that each individual letter corresponds to one single sound, they will interact with, use and manipulate the phonemic level in different ways (to spell, for rhymes, puns and similar language play, etc.). Speakers may even come to think of sounds in terms of letters, since the latter provide a visual representation of the cluster of articulatory parameters that the production of the members of phoneme categories involve.

The structural salience of the system is no doubt reinforced for subjects trained in alphabetic writing but perhaps not exclusively caused by it. Before learning an alphabetic writing system (or any other type of writing system), speakers already posses some *phonological awareness* or degree of sensitivity to the sound structure (mainly of word structure) of oral language like syllables, onsets, rhymes, or phonemes as shown by their ability to recognise, discriminate, and manipulate the sounds of the language (see e.g. Bryant, 1990; Goswami & Bryant, 1990). In this respect, research on phonological awareness has shown that, irrespective of their language background, children become increasingly sensitive to smaller and smaller parts of words as they grow older. Children can detect or manipulate

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

syllables before they can detect or manipulate onsets and rimes, and they can detect or manipulate these before they can detect or manipulate individual phonemes within intrasyllabic word units (Anthony & Francis, 2005: 256). As far as phoneme awareness is concerned, research has also shown that characteristics of oral languages (e.g. saliency and complexity of word structures, phoneme positions, articulatory factors, etc.) influence the degree of awareness of phonemes among pre-literate children (Anthony & Francis, 2005: 257), although "…most children achieve minimal levels of phoneme awareness prior to literacy instruction" (ibid) and the same is true of adult illiterates (see e.g. Adrian *et al.*, 1995; Loureiro *et al*. 2004; Tarone & Bigelow, 2005 for a review). However, phoneme-level awareness and skills develop fast once alphabetic writing is learned -and faster in children acquiring orthographically consistent languages with consistent spelling-to-sound and consistent sound-to-spelling relations like Italian or German (Anthony & Francis, 2005: 257; Goswami, 2002). In any case, the relationship between literacy and developing phonological awareness (including phoneme awareness) appears to be reciprocal (Anthony & Francis, 2005; Perfetti *et al*. 1987). Children's preliterate phonological awareness and the phonological awareness they develop while learning the names and sounds of letters in their alphabet help children learn to read but reading and writing provide feedback that influences individual's phonological awareness development. According to Ravid and Tolchinsky (2002: 432), "….specific aspects of language awareness, especially phonological and morphological awareness, both promote and are promoted by learning to read and write...".

VI. CONCLUSION

This paper has provided some empirical evidence on the potential existence of a basic level in taxonomies of phonological categories that are plausible for adult literate speakers of English. In this respect, four CF experiments were carried out to find out whether any of the four categories studied at different levels of abstraction in the taxonomy, i.e. CONSONANT, PLOSIVE, PHONEME P, and ALLOPHONE P, was more salient as shown by the ease with which phonetically naïve subjects could 'form' the categories. The results show that the category PHONEME P was the easiest to form, suggesting that the phoneme level may have some sort of basic-level status in phonological taxonomies. The reasons why this level could be more salient were also discussed and it was claimed that they might be due to structural (e.g. greater similarity of the members of the category, i.e. its allophones and greater distinctiveness from

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

other categories) as well as cultural factors (e.g. greater salience boosted by alphabetic literacy).

Suggestive as the results obtained are, they are limited and more evidence should be obtained to confirm the finding that the phoneme level is the basic level of phonological taxonomies for phonetically naïve subjects. Directions for future research could also involve looking at what the basic level is in taxonomies for subjects whose language uses a writing system that is not alphabetic, which may reveal that phoneme might not be the basic level of abstraction for phonological taxonomies from a universal point of view (but rather depend strongly on the alphabetic/non-alphabetic spelling of the subject's language).

NOTES

1. I express my thanks to the numerous people who have provided their input on this paper and/or the work involved in it and reported in Mompeán (2002). These include, amongst others, Helen Fraser, René Dirven, David Eddington, Antonio Barcelona, Lorenzo Fernández Maimó and Pilar Mompeán.

2 These include not only natural or artefactual categories (e.g. Rosch *et al.*, 1976; Rosch, 1978) but also artificial categories (Mervis & Crisafi, 1982; Murphy & Smith, 1982) and a host of other types of categories like environmental scenes (Tversky, 1986, 1990; Tversky & Hemenway, 1983, 1984), events (Morris & Murphy, 1990; Rifkin, 1985), social, ideological, cultural and psychological situations (Cantor *et al.*, 1982), psychiatric diagnoses (Cantor *et al.*, 1980), traits (Brewer *et al.*, 1981; Cantor & Mischel, 1979; Dahlgren, 1985; John *et al.*, 1991), emotions (Fehr & Russel, 1984; Shaver *et al.*, 1987), computer programming concepts (Adelson, 1985), sentences (Corrigan, 1991), etc.

3 On the contrary, fewer attributes are listed for category members at the superordinate level (e.g. FURNITURE, VEHICLE, ANIMAL) and there is virtually no increase for subordinate categories (e.g. ROCKING CHAIR, SPORTS CAR, RETRIEVER) over the basic level unless expert knowledge is developed for them (e.g. Tanaka & Taylor 1991). On a related note, it has also been found that at least for natural and artefactual categories, most of the attributes listed for both basic-level and subordinate categories refer to physical parts like "arms", "legs", "eyes", etc. (e.g. Hemenway, 1981; Mervis & Greco, 1984; Tversky, 1986, 1990, 1991; Tversky & Hemenway, 1983, 1984, 1991). However, parts are neither necessary nor sufficient for establishing a basic-level structure (Murphy, 1991a, 1991b). The few attributes listed for superordinate categories are abstract attributes that refer to the functions of objects.

4. In contrast, members of superordinate categories like FURNITURE do not share a common shape and, as a consequence, a calculated average shape of a number of superordinate objects is not readily recognisable as a member of that superordinate category. Also, some gain in similarity of shapes occurs for subordinate category members (e.g. different instances of the category KITCHEN CHAIR) but this increase in similarity is so small when going from the basic to the subordinate level that the basic level is again preferred (Rosch *et al.* 1976).

5. In contrast, fewer similar motor actions are used to interact with members of superordinate categories. Also, although subjects behave in a very similar way with members of subordinate categories, no more movements are made in common to subordinate than to basic-level categories (Rosch, 1978; Rosch *et al.*, 1976)

6. This is so unless the to-be-identified object (e.g. a chair) has to be categorised as part of a scene or context, such as living room with a sofa, tables, and lamps in which case categorising the object at the superordinate level (e.g. "furniture") is just as fast as categorising the object at the basic level, for example "chair" (Murphy & Wisniewski, 1989a) or when subjects possess a high degree of expertise, in which case spontaneous naming of entities occurs at the subordinate level (Tanaka & Taylor, 1991). In general, the need for specificity or generality

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

in the information conveyed may require the use of subordinate or superordinate level category names (Cruse, 1977; Rosch *et al.*, 1976).

7. Several studies (e.g. Lassaline *et al.*, 1992; Mervis & Crisafi, 1982; Murphy & Smith, 1982; Murphy & Brownell 1985) rule out the possibility that the basic level is due to linguistic factors like word length and frequency, which reflect properties of the category names rather than properties of conceptual representations. However, basic-level names differ from category names of superordinate and subordinate categories since they are typically shorter, underived, morphosyntactically regular, etc. (see Berlin, 1978, 1992; Brown, 1958, 1976; Brown *et al.*, 1976; Mervis & Rosch, 1981 for lengthier discussions), or the first to be learned developmentally (Anglin, 1977; Blewitt & Durkin, 1982; Dougherty, 1978; Mervis, 1980, 1984; Mervis & Mervis, 1982; Poulin-Dubois *et al.*, 1995; Rescolda, 1980; Shipley *et al.*, 1983; Stross, 1973; White, 1982) or primarily used by parents or caretakers in their speech to children (Anglin. 1977; Blewitt, 1983; Brown, 1956, 1976; Callanan, 1983, 1985; Poulin-Dubois *et al.*, 1995; Shipley *et al.*, 1993; White, 1982).

8. However, despite the common belief that first-learned words correspond with first-learned categories (both described as 'basic level') leading to the belief that language acquisition is a reasonably good indicator of early cognition, this is not necessarily so since toddlers, for instance, often overextend their first words. McDonough (2002), for instance, conducted two experiments that examined two-year-olds' production and comprehension of basic-level terms. The results showed overextensions both in production (e.g. children labelled a rocket 'airplane') and comprehension (e.g. they pointed to a rocket when airplane was requested). McDonough argues that toddlers extend labels to a wider conceptual domain because they have not clearly differentiated basic-level concepts from related conceptual categories.

9. In any case, even in partonomic organisation of phonological categories we can look into questions of what the basic level is, since the existence of a basic level has also been claimed for partonomies (see e.g. Rifkin 1985; Tversky 1989, 1990) but explicit reference should be made to whether taxonomic or partonomic organisation is being discussed. This issue can be linked to discussions in the speech perception literature on the 'basic' unit of speech perception (see e.g. Goldinger & Azuma, 2003), phonological awareness and phonemic awareness literature (e.g. Read *et al.* 1986), or the 'basic-level' salience of the phoneme, the syllable or the phonological word for speakers with different writing systems (alphabetic, logographic, syllabary-based, etc.)

10. The CF paradigm was originally and extensively used in psychology during the behavioural and information processing eras for a wide range of purposes. The technique has also been employed to address different phonological and/or phonetic questions (e.g. Jaeger, 1980, 1984; Jaeger & Ohala, 1984; Wang & Derwing, 1986; Weitzman, 1992).

11. As pointed out by Taylor (2006), the word *category formation* can be conceived of as a problem solving task in which subjects have to work out the criteria by which a given set of stimuli have been put into a certain category while other stimuli have not. Thus, following Kendler (1961: 447), the noun *concept formation* should be understood merely as referring to a well-known experimental technique, not an abstract psychological process.

12. A further aim of the test session of a CF experiment is to find out about the way the subject classifies instances whose category membership is controversial or unclear for some reason. These stimuli are called *test* tokens and they provide the experimenter with information about the boundaries of categories previously 'formed' by the subject during the training session. Test tokens were included in the first three experiments since the latter also looked at other phonological problems investigated in earlier work by the author (Mompeán, 2002). These problems included the assignment of the so-called semi-vowels in English (i.e. /w, j/) to the category CONSONANT or not, the assignment of English affricates (i.e. /ʧ, ʤ/) to the category PLOSIVE, or the treatment of plosives after tautosyllabic /s/ as instances of the fortis (voiceless) plosives (i.e. /p, t, k/) or not.

13. However, as Mompeán (2004) argues, the members of one and the same phoneme category need not share all the features in common and there may be no single feature that is shared by all members of the category.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

REFERENCES

Abercrombie, D. (1967). *Elements of general phonetics.* Edinburgh: Edinburgh University Press.

Adelson, B. (1985). Comparing natural and abstract categories: A case study from computer science. *Cognitive Science*, 9, 417-430.

Adrian, J. A., Alegría, J., & Morais, J. (1995). Metaphonological abilities of Spanish illiterate adults. *International Journal of Psychology*, 30, 329–353.

Anglin, J. M. (1977). *Word, object, and conceptual development.* New York: W. W. Norton & Company.

Anthony, J. L., & Francis, D. J. 2005. Development of Phonological Awareness. *Current Directions in Psychological Science,* 14(5):255-259.

Berlin, B. (1978). Ethnobiological classification. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and categorisation.* Hillsdale, NJ: LEA, pp. 9-26.

Berlin, B. (1992). *Ethnobiological classification: Principles of categorisation of plants and animals in traditional societies*. Princetown, NJ: Princetown University Press.

Berlin, B., Breedlove, D. E., & Raven, P. H. (1973). General principles of classification and nomenclature in folk biology. *American Anthropologist*, 75, 214-242.

Blewitt, P. (1983). Dog vs. collie: Vocabulary in speech to young children. *Developmental Psychology*, 19, 602-609.

Blewitt, P. (1994). Understanding categorical hierarchies: The earliest levels of skill. *Child Development*, 65, 1279-1298.

Blewitt, P., & Durkin, M. (1982). Age, typicality, and task effects on categorization of objects. *Perception and Motor Skills*, 55, 435-445.

Bolton, N. (1977). *Concept formation*. Oxford: Pergamon Press.

Bourne, L. E. (1966). *Human conceptual behavior*. Boston, NJ: Allyn & Bacon.

Bourne, L. E. (1970). Knowing and using concepts. *Psychological Review*, 77, 546-556.

Brewer, M. B., Dull, V., & Lui, L. (1981). Perceptions of the elderly: Stereotypes as prototypes. *Journal of Personality and Social Psychology*, 41, 656-670.

Brown, C. H., Kolar, J., Torrey, B. J., Truong-Quang, T., & Volkman, P. (1976). Some general principles of biological and non-biological folk classification. *American Ethnologist*, 3, 73-85.

Brown, R. W. (1958). How shall a thing be called?. *Psychological Review*, 65, 14-21.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Brown, R. W. (1976). Reference: In memorial tribute to Eric Lenneberg. *Cognition*, 4, 125-153.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: John Wiley & Sons.

Bryant, P. (1990). Phonological development and reading. In P. D. Pumfrey, & C. D. Elliot (Eds.), Children's difficulties in reading, spelling and writing Challenges and responses. London The Falmer Press, pp. 63-82.

Callanan, M. A. (1983). *Parental input and young children's acquisition of hierarchically organized concepts.* Ph.D. Diss. Stanford University, Palo Alto, CA.

Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development*, 56, 508-523.

Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. *Journal of Personality and Social Psychology*, 35: 38-48.

Cantor, N., Mischel, W., & Schwartz, J. C. (1982). A prototype analysis of psychological Situations. *Cognitive Psychology*, 14, 45-77.

Cantor, N., Smith, E. E., French, R. D., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorisation. *Journal of Abnormal Psychology*, 89, 181-193.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.

Corrigan, R. (1991). Sentences as categories: Is there a basic-level sentence?. *Cognitive Linguistics*, 2, 339-356.

Corter, J. E., & Gluck, M. A. (1992). Explaining basic categories: Feature predictability and information. *Psychological Bulletin*, 111, 291-303.

Cruse, D. A. (1977). The pragmatics of lexical specificity. *Journal of Linguistics*, 13, 153-164.

Dahlgren, K. (1985). The cognitive structure of social categories. *Cognitive Science*, 9, 379-398.

Dell, G. S., & Newman, J. E. (1980). Detecting phonemes in fluent speech. *Journal of Verbal Learning and Verbal Behavior*, 19, 608-623.

Donegan, P. J., & Stampe, D. (1979). The study of natural phonology. In D. A. Dinnsen (Ed.), *Current approaches to phonological theory*. Bloomington: Indiana University Press, pp. 126-173.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Dougherty, J. W. D. (1978). Salience and relativity in classification. *American Ethnologist*, 5, 66-80.

Downing, P. (1977). On 'basic levels' and the categorisation of objects in English discourse. *Berkeley Linguistics Society*, 3, 475-87.

Fehr, B. & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, 113, 464-486.

Fodor, J. A., Garrett, M. F., & Brill, S. L. (1975). Pi-ka-pu: The perception of speech sounds by pre-linguistic infants. *Perception and Psychophysics*, 18, 74-78.

Fraser, H. (2006). Phonological concepts and concept formation: Metatheory, theory and application. *International Journal of English Studies*, 6(2), 55-75.

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: categories and inferences in 2-year-old children. *Developmental Psychology*, 26, 796-804.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23, 183-209.

Gelman, S. A., & O'Reilly, A. W. (1988). Children's inductive inferences within superordinate categories: The role of language and category structure. *Child Development*, 59, 876-887.

Goldinger, S. D., & Azuma, T. 2003. Puzzle-solving science: The quixotic quest for units in speech perception. *Journal of Phonetics*, 31, 305-320.

Goswami, U. 2002. In the beginning was the rhyme? A reflection on Hulme, Hatcher, Nation, Brown, Adams and Stuart, 2001. *Journal of Experimental Child Psychology, 82*, 47-57.

Goswami, U., & Bryant, P. E. (1990). *Phonological skills and learning to read*. London: Erlbaum.

Hampton, J. A. (1982). A demonstration of intransitivity in natural concepts. *Cognition*, 12, 151-164.

Hampton, J. A. (1988). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learing, Memory, and Cognition*, 14, 12-32.

Hemenway, K. (1981). *The role of perceived parts in categorisation*. Ph.D. Diss. Stanford University.

Honeck, R. P., Firment, M., & Case, T. J. (1987). Expertise and categorisation. *Bulletin of the Psychonomic Society*, 25, 431-434.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Horton, M. S., & Markman, E. M. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development*, 51, 708-719.

Hull, C. L. (1920). Quantitative aspects of the evolution of concepts: an experimental study. *Psychological Monographs,* 28. No. 1. (Whole No. 123).

Hunt, E. B. (1962). *Concept learning: An information processing problem.* New York: John Wiley & Sons.

Jaeger, J. J. (1980). *Categorisation in phonology: An experimental approach*. Ph.D. diss., University of California, Berkeley.

Jaeger, J. J. (1984). Assessing the psychological status of the Vowel Shift rule. *Journal of Psycholinguistic Research*, 13, 13-56.

Jaeger, J. J. (1986). Concept formation as a tool for linguistic research. In J. J. Ohala & J. J. (eds.). 1986. *Experimental phonology*. Orlando: Academic Press, pp. 211-238.

Jaeger, J. J., & Ohala, J. J. (1984). On the structure of phonetic categories. *Berkeley Linguistics Society*, 10, 15-26.

John, O. P., Hampson, S. E., & Goldberg, L. R. (1991). The basic level in personality-trait hierarchies: studies of trait use and accessibility in different contexts. *Journal of Personality and Social Psychology,* 60, 348-361.

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, 16, 243-275.

Jones, G. V. (1983). Identifying basic categories. *Psychological Bulletin*, 94, 423-428.

Katz, J. J., & Postal, P. M. (1964). *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.

Kendler, T. S. (1961). Concept formation. *Annual Review of Psychology*, 12, 447-472.

Kreidler, C. W. (1989). *The pronunciation of English: A course book in phonology.* Oxford, UK: Blackwell.

Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Lassaline, M. E., Wisniewski, E. J., & Medin, D. L. (1992). The basic level in artificial and natural categories: Are all levels created equal?. In B. Burns (Ed.), *Percepts, concepts, and categories: The representation and processing of information*. Amsterdam: Elsevier, pp. 327-278.

Lin, E. L., Murphy, G. L., & Shoben, E. J. (1997). The effects of prior processing episodes on basic-level superiority. *Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology*, 50: 25-48.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Loureiro, C de Santos, Braga, W. L., do Nascimento Souza, L., Nunes Filho, G., Queiroz, E. & Dellatolas, G. (2004). Degree of illiteracy and phonological and metaphonological skills in unschooled adults. *Brain and Language*, 89(3), 499-502.

Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition*, 24, 31–45.

Margolis, E. (1994). A reassessment of the shift from the classical theory of concepts to prototype theory. *Cognition*, 51, 73-89.

Markman, E. M., Horton, M. S., & McLanahan, A. G. (1980). Classes and collections: Principles of organization in the learning of hierarchical relations. *Cognition*, 8, 227-241.

McCloskey, M. E. (1980). The stimulus familiarity problem in semantic memory research. *Journal of Verbal Learning and Verbal Behavior*, 19, 485-502.

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well-defined or fuzzy sets?. *Memory and Cognition*, 6, 462-472.

McCloskey, M. E., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic Memory. *Cognitive Psychology*, 11, 1-37.

McDonough, L. (2002). Basic-level nouns: First learned but misunderstood. *Journal of Child Language*, 29(2), 357-377.

Medin, D. L. (1983). Structural Principles in Categorisation. In T. J. Tighe & B. E. Shepp (Eds.), *Perception, cognition, and development: Interactional analyses.* Hillsdale, NJ: LEA, pp. 203-230.

Medin, D. L., Lynch, E. B., & Coley, J. D. (1997). Categorisation and reasoning among tree experts: Do all roads lead to Rome?. *Cognitive Psychology*, 32, 49-96.

Mervis, C. B. (1980). Category structure and the development of categorisation. In R. J. Spiro, B. C. Bruce & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: LEA, pp. 279-307.

Mervis, C. B. (1984). Early lexical development: the contributions of mother and child. In C. Sophian (Ed.), *Origins of cognitive skills*. Hillsdale, NJ: LEA, pp. 339-370.

Mervis, C. B. (1987). Child-basic object categories and early lexical development. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorisation.* New York: CUP, pp. 201-233.

Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic- and superordinate-level categories. *Child Development*, 53, 258-266.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Mervis, C. B., & Greco, C. (1984). Parts and early conceptual development: Comment on Tversky and Hemenway. *Journal of Experimental Psychology: General*, 113, 194-197.

Mervis, C. B., & Mervis, C. A. (1982). Leopards are kitty-cats: Object labeling by mothers for their thirteen-month-olds. *Child Development*, 53, 267-273.

Mervis, C. B., & Rosch, E. (1981). Categorisation of natural objects. *Annual Review of Psychology*, 32, 89-115.

Mompeán, J. A. (1999). A cognitive view of the concept of the phoneme. *6th International Cognitive Linguistics Conference*. Stockholm University, July 1999.

Mompeán, J. A. (2002). *The categorisation of the sounds of English: Experimental evidence in phonology.* Unpublished Ph.D dissertation. University of Murcia.

Mompeán, J. A. (2004). Category overlap and neutralisation: The importance of speakers' classification in phonology. *Cognitive Linguistics*, 15, 429-469

Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. *Memory and Cognition*, 18, 407-418.

Murphy, G. L. (1991a). Parts in object concepts: Experiments with artificial categories. *Memory and Cognition*, 19, 423-438.

Murphy, G. L. (1991b). More on parts in object concepts: Response to Tversky and Hemenway. *Memory and Cognition*, 19, 443-447.

Murphy, G. L., & Brownell, H. H. (1985). Category Differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 70-84.

Murphy, G. L, & Lassaline, M. E. (1997). Hierarchical structure in concepts and the basic level of categorisation. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts, and categories*. Cambridge, MA: MIT Press, pp. 93-131.

Murphy, G. L., & Smith, E. E. (1982). Basic-level superiority in picture categorisation. *Journal of Verbal Learning and Verbal Behavior*, 21, 1-20.

Murphy, G. L., & Wisniewski, E. J. (1989a). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 572-586.

Murphy, G. L., & Wisniewski, E. J. (1989b). Feature correlations in conceptual representations. In G. Tiberghien (Ed.), *Advances in cognitive science. Vol. 2: theory and applications*. Chichester, England: Ellis Horwood, pp. 23-45.

Murphy, G. L., & Wright, J. C. (1984). Changes in conceptual structure with expertise: Differences between real-world experts and novices. *Journal of Experimental Psycholgy: Learning, Memory, and Cognition*, 10, 144-155.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

O'Connor, J. D. (1973). *Phonetics.* Harmondsworth, Penguin Books.

Nathan, G. S. (1986). Phonemes as mental categories. *Berkeley Linguistics Society*, 12, 212-223.

Nathan, G. S. (1996). Steps towards a cognitive phonology. In B. Hurch & R. Rhodes (Eds.), *Natural Phonology: The state of the art*. Berlin: Mouton de Gruyter, pp. 107-120.

Nathan, G. S. (1999). What functionalists can learn from formalists in phonology. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan, & K. Wheatley (Eds.), *Functionalism and formalism in linguistics*. Vol. 1: *General papers*. Amsterdam/Philadelphia: John Benjamins, pp. 305-327.

Nathan, G. S. (2007). *Phonology in Cognitive Grammar*. Manuscript.

Neisser, U., & Weene, P. (1962). Hierarchies in concept attainment. *Journal of Experimental Psychology*, 64, 640-645.

Newman, J. E., & Dell, G. S. (1978). The phonological nature of phoneme monitoring: A critique of some ambiguity studies. *Journal of Verbal Learning and Verbal Behavior*, 17, 359-374.

Ohala, J. J. (1986). Consumer's guide to evidence in phonology. *Phonology Yearbook*, 3, 3-26.

Pansky, A., & Koriat, A. 2004. The basic-level convergence effect in memory distortions. *Psychological Science*, 15(1), 52-59.

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. 1987. Phonemic knowledge and learning to read are reciprocal. *Merrill-Palmer Quarterly*, 33, 283-319.

Pike, K. L. (1943). *Phonetics: A critical analysis of theory and a technic for the practical description of sounds*. Ann Arbor, MI: The University of Michigan Press.

Poulin-Dubois, D., Graham, S. & Sippola, L. (1995). Early lexical development: The contribution of parental labelling and infants' categorisation abilities. *Journal of Child Language*, 22, 325-343.

Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language*, 29, 417–447.

Rescolda, L. A. (1980). Overextension in early language development. *Journal of Child Language*, 7, 321-335.

Rey, G. (1983). Concepts and stereotypes. *Cognition*, 15, 237-262.

Rey, G. (1985). Concepts and conceptions: a reply to Smith, Medin, and Rips. *Cognition*, 19, 297-303.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Rifkin, A. (1985). Evidence for a basic level in event taxonomies. *Memory and Cognition*, 13, 538-556.

Rosch, E. H. (1978). Principles of categorisation. In E. H. Rosch & B, B. Lloyd (Eds.), *Cognition and categorisation*. Hillsdale, NJ: LEA, pp. 27-48.

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.

Rosch, E. H., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52, 1061-1086.

Shipley, E. F. (1993). Categories, hierarchies, and induction. In D. L. Medin (Ed.), *The psychology of learning and motivation.* Vol. 30. San Diego: Academic Press, pp. 265-301.

Shipley, E. F., Kuhn, I. F., & Madden, E. C. (1983). Mothers' use of superordinate category terms. *Journal of Child Language*, 10, 571-588.

Sloman, S. A. (1997). Categorical inference is not a tree: The myth of inheritance hierarchies. *Cognitive Psychology*, 35, 1-13.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 81, 214-241.

Smith, E. E., Balzano, G. J., & Walker, J. (1978). Nominal, perceptual, and semantic codes in picture categorisation. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic factors in cognition*. Hillsdale, NJ: LEA, pp. 137-168.

Stemberger, J. P., Elman, J. L., & Haden, P. (1985). Interference between phonemes during phoneme monitoring: Evidence for an interactive activation model of speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 475-489.

Stross, B. (1973). Acquisition of botanical terminology by Tzeltal children. In M. S. Edmonson (Ed.). *Meaning in Mayan languages.* The Hague: Mouton, pp. 107-141.

Sutcliffe, J. P. (1993). Concept, class, and category in the tradition of Aristotle. In I. van Mechelen, J. Hampton, R. S. Michalski & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data*. New York: Academic Press, pp. 35-65.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: is the basic level in the eye of the beholder?. *Cognitive Psychology*, 23, 457-482.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Tarone, E., & Bigelow, M. (2005). Impact of literacy on oral language processing: Implications for second language acquisition research. *Annual Review of Applied Linguistics*, 25, 77–97.

Taylor, J. R. (2002). *Cognitive grammar*. Oxford: Oxford University Press.

Taylor, J. R. (2003). *Linguistic categorization*. Oxford: Oxford University Press. First edition: 1989.

Taylor, J. R. (2006). Where do phonemes come from? A view from the bottom. *International Journal of English Studies*, 6(2), 19-54.

Tversky, B. (1985). Development of taxonomic organization of named and pictured categories. *Developmental Psychology*, 21, 1111-1119.

Tversky, B. (1986). Components and categorisation. In C. Craig (Ed.), *Noun classes and categorisation*. Amsterdam: John Benjamins, pp. 63-75.

Tversky, B. (1989). Parts, partonomies, and taxonomies. *Developmental Psychology*, 25, 983-995.

Tversky, B. (1990). Where partonomies and taxonomies meet. In S. L. Tsohatzidis (Ed.), *Meanings and prototypes: Studies in linguistic categorisation*. London: Routledge, pp. 334-344.

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology*, 15, 121-149.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, 113, 169-193.

Tversky, B., & Hemenway, K. (1991). Parts and the basic level in natural categories and artificial stimuli: Comments on Murphy (1991). *Memory and Cognition*, 19, 439-442.

Ungerer, F. (1994). Basic level concepts and parasitic categorisation: A cognitive alternative to conventional semantic hierarchies. *Zeitschrift für Anglistik und Amerikanistik*, 42: 148-162.

Ungerer, F. & Schmid, H. J. (1996) *An introduction to cognitive linguistics*. New York: Addison Wesley Longman.

URL1 *Wikipedia* entry 'Linnean taxonomy'. http://en.wikipedia.org/wiki/Linnaean_taxonomy

Wang, H. S., & Derwing, B. L. (1986). More on English Vowel shift: The back vowel question. *Phonology Yearbook*, 3, 99-116.

Waxman, S. R., Shipley, E. F., & Shepperson, B. (1991). Establishing new subcategories: The role of category labels and existing knowledge. *Child Development*, 62, 127-138.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Weitzman, R. (1992). Vowel categorisation and the critical band. *Language and Speech*, 35, 115-125.

Weitzman, R. (1993). How to get the horse to open its mouth: Using the concept formation paradigm in speech perception research. In J. A. Nevis *et al.* (Eds.), *Papers in honor of Frederick Brengelman on the occasion of the 25th anniversary of the Linguistics Department at California State University.* Fresno, CA: California State University, pp. 141-149.

White, T. G. (1982). Naming practices, typicality, and underextension in child language. *Journal of Experimental Child Psychology*, 33, 324-346.

Wisniewski, E. J., Imai, M., & Casey, L. (1996). On the equivalence of superordinate concepts. *Cognition*, 60, 269-298.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

APPENDIX

Stimulus List for the category CONSONANT (exp. 1)						
Ord.	Item	P(+) N(-)		Ord.	Item	P(+) N(-)
LEARNING SESSION						
1.	path	+		53.	change	+
2.	ash	-		54.	aim	-
3.	boom	+		55.	jug	+
4.	ache	-		56.	ebb	-
5.	toad	+		57.	ace	-
6.	duck	+		58.	mug	+
7.	kid	+		59.	itch	-
8.	up	-		60.	name	+
9.	give	+				
10.	eat	-		**TEST SESSION**		
11.	seethe	+				
12.	zone	+				
13.	edge	-		1.	pet	+
14.	fish	+		2.	bath	+
15.	van	+		3.	tooth	+
16.	at	-		4.	ode	- c
17.	egg	-		5.	deep	+
18.	ill	-		6.	hot	test
19.	thing	+		7.	eight	-
20.	off	-		8.	ape	-
21.	that	+		9.	cab	+
22.	each	-		10.	guess	+
23.	hours	-i		11.	oath	- c
24.	cheese	+		12.	fetch	+
25.	job	+		13.	heat	test
26.	miss	+		14.	us	-
27.	out	-		15.	vague	+
28.	neck	+		16.	all	- c
29.	eve	-		17.	seed	+
30.	on	-		18.	youth	test
31.	pub	+		19.	ship	+ c
32.	aid	-		20.	wit	test
33.	beach	+		21.	zoom	+
34.	teach	+		22.	earn	-
35.	oil	-		23.	thin	+
36.	odd	-		24.	then	+
37.	dove	+		25.	urge	-
38.	arm	-		26.	shock	+ c
39.	call	+		27.	orb	- c
40.	goose	+		28.	once	test
41.	ale	-		29.	chase	+
42.	safe	+		30.	yet	test
43.	earth	-		31.	of	-
44.	zip	+		32.	judge	+
45.	fang	+		33.	wall	test
46.	owl	-		34.	if	-
47.	ooze	-		35.	hard	test
48.	ice	-		36.	shell	+ c
49.	vet	+		37.	map	+
50.	thick	+		38.	net	+
51.	age	-		39.	use	test
52.	those	+		40.	art	-

Stimulus List for the Category PHONEME P (exp. 3)						
Ord.	Item	P(+) N(-)		Ord.	Item	P(+) N(-)
LEARNING SESSION						
1.	pet	+		53.	print	+
2.	sell	-		54.	fee	-
3.	up	+		55.	pond	+
4.	egg	-		56.	end	-
5.	pay	+		57.	grant	-
6.	plea	+		58.	top	+
7.	drip	+		59.	fist	-
8.	die	-		60.	trap	+
9.	apt	+				
10.	tray	-		**TEST SESSION**		
11.	priest	+				
12.	depth	+				
13.	drill	-		1.	pit	+
14.	path	+		2.	pear	+
15.	ape	+		3.	prow	+
16.	old	-		4.	sheet	-
17.	drift	-		5.	plane	+
18.	golf	-		6.	spend	test
19.	pie	+		7.	near	-
20.	fish	-		8.	slow	-
21.	pray	+		9.	clamp	+
22.	ash	-		10.	pulse	+
23.	bay	-i		11.	bear	- i
24.	place	+		12.	cap	+
25.	opt	+		13.	spa	test
26.	stamp	+		14.	ground	-
27.	sphere	-i		15.	prayer	+
28.	post	+		16.	false	-
29.	graph	-i		17.	drop	+
30.	blast	-i		18.	spy	test
31.	shop	+		19.	glimpse	+ c
32.	east	-		20.	spoon	test
33.	pea-p	+		21.	prince	+
34.	play	+		22.	phone	- i
35.	self	-		23.	paste	+
36.	psalm	-i		24.	ship	+
37.	proud	+		25.	sly	-
38.	sea-see	-		26.	lapsed	+ c
39.	asp	+		27.	slob	-c, i
40.	clasp	+		28.	sponge	test
41.	dry	-		29.	plot	+
42.	damp	+		30.	spray	test
43.	clean	-		31.	cross	-
44.	keep	+		32.	tramp	+
45.	paw	+		33.	sply	test
46.	act	-		34.	nymph	-i
47.	bet	-i		35.	spring	test
48.	trust	-		36.	pure	+ c
49.	plough	+		37.	lamp	+
50.	group	+		38.	rapt	+
51.	fond	-		39.	split	test
52.	imp	+		40.	stealth	-

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Stimulus List for the category PLOSIVE (exp. 2)						
Ord.	Item	P(+) N(-)		Ord.	Item	P(+) N(-)
LEARNING SESSION						
1.	push	+		53.	cave	+
2.	fall	-		54.	zone	-
3.	bus	+		55.	gull	+
4.	verse	-		56.	shell	-
5.	tall	+		57.	thighs	- int
6.	cash	+		58.	path	+
7.	gas	+		59.	thus	- int
8.	safe	-		60.	ton	+
9.	pace	+				
10.	zeal	-		**TEST SESSION**		
11.	beef	+				
12.	tough	+				
13.	shove	-		1.	pave	+
14.	kill	+		2.	bath	+
15.	girl	+		3.	tool	+
16.	thief	- i		4.	file	- c
17.	these	- i		5.	coal	+
18.	fish	-		6.	chill	test
19.	pill	+		7.	seethe	-
20.	verve	-		8.	veil	-
21.	bill	+		9.	goose	+
22.	sell	-		10.	pause	+
23.	zoos	-		11.	zoom	-
24.	toes	+		12.	ball	+
25.	case	+		13.	jazz	test
26.	gaze	+		14.	shoal	-
27.	knife	- i		15.	dish	+ c
28.	pile	+		16.	thieve	- i
29.	psalm	- i		17.	tale	+
30.	kneel	- i		18.	chief	test
31.	beige	+		19.	deaf	+ c
32.	shave	-		20.	juice	test
33.	tail	+		21.	call	+
34.	cough	+		22.	those	- i
35.	thaws	- i		23.	give	+
36.	miss	- i		24.	pull	+
37.	goal	+		25.	moth	- int
38.	nerve	- i		26.	dull	+ c
39.	puff	+		27.	nose	- i
40.	booze	+		28.	choose	test
41.	this	- i		29.	buzz	+
42.	tease	+		30.	cheese	test
43.	mill	- i		31.	fill	-
44.	cool	+		32.	time	+
45.	gash	+		33.	jail	test
46.	five	-		34.	gnash	- c
47.	mass	- i		35.	choice	test
48.	veal	-		36.	deal	+ c
49.	pass	+		37.	cause	+
50.	bush	+		38.	guess	+
51.	sauce	-		39.	jaws	test
52.	toll	+		40.	these	- i

Stimulus List for the Category ASPIRATED P (exp. 4)						
Ord.	Item	P(+) N(-)		Ord.	Item	P(+) N(-)
LEARNING SESSION						
1.	paw	+		53.	poise	+
2.	spend	-		54.	drop	-
3.	push	+		55.	pan	+
4.	rapt	-		56.	spy	-
5.	post	+		57.	clasp	-
6.	pray	+		58.	pots	+
7.	power	+		59.	ship	-
8.	stamp	-		60.	paled	+
9.	pass	+				
10.	up	-		**TEST SESSION**		
11.	pulse	+				
12.	plot	+				
13.	spray	-		1.	payer	+
14.	purr	+		2.	pass	+
15.	pin	+		3.	pegs	+
16.	spoon	-		4.	span	-
17.	spear	-		5.	pen	+
18.	damp	-		6.	clamp	-
19.	pond	+		7.	wept	-
20.	depth	-		8.	top	-
21.	proud	+		9.	pines	+
22.	help	-		10.	pea	+
23.	shop	-		11.	spa	-
24.	pay	+		12.	pool	+
25.	pill	+		13.	palm	+
26.	pence	+		14.	camp	-
27.	cap	-		15.	pie	+
28.	play	+		16.	glimpse	- c
29.	split	-		17.	peace	+
30.	apt	-		18.	keep	-
31.	par	+		19.	poles	+
32.	spring	-		20.	spill	-
33.	pain	+		21.	pear	+
34.	piles	+		22.	tramp	-
35.	spice	-		23.	pun	+
36.	caps	-		24.	pubs	+
37.	prince	+		25.	lapsed	-
38.	gulp	-		26.	pure	+ c
39.	pace	+		27.	group	-
40.	pelt	+		28.	pause	+
41.	trap	-		29.	paved	+
42.	plea	+		30.	sponge	-
43.	splay	-		31.	gasp	-
44.	paste	+		32.	peel	+
45.	print	+		33.	opt	-
46.	spare	-		34.	ape	-
47.	kept	-		35.	punch	+
48.	lamp	-		36.	pull	+
49.	plough	+		37.	pave	+
50.	priest	+		38.	pant	+
51.	ropes	-		39.	drip	-
52.	puff	+		40.	gasp	-

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Is the Phoneme Usage-Based? – Some Issues[1]

GEOFFREY S. NATHAN[*]
Wayne State University

ABSTRACT

After a brief review of the history of the phoneme, from its origins in the nineteenth century to Optimality Theory, including some Cognitive Linguists' views of the concept. I argue that current 'usage-based' theorists views of the phoneme may not be able to explain some facts about how naïve speakers process language, both consciously and subconsciously. These facts include the invention of and worldwide preference for alphabetic writing systems, and language processing evidence provided by Spoonerisms, historical sound changes affecting all (or most) lexical items in a language and each other, and the fact that allophonic processes normally do not show lexical conditioning. I further suggest that storing speech in terms of a small number of production/perception units such as phonemes could be due to the fact that phonemes seem to optimize both efficiency and informativeness in much the same way as other basic-level categories.

KEYWORDS: history of phonology, phonemic processing, usage-based theories, natural phonology

[*] ***Address for correspondence***: Geoffrey S. Nathan, Department of English, Wayne State University, 5057 Woodward, Detroit, MI, 48202, USA. Tel. 1 313 577-8621, Fax: 1 313 577-0404, e-mail geoffnathan@wayne.edu

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I. INTRODUCTION AND BRIEF HISTORY OF THE PHONEME

The phoneme has had a checkered history—it was first proposed as a name for something like a speech sound by Dufriche-Desgenettes (Anderson, 1985), but soon evolved into essentially its current sense of a minimal unit of sound in which lexical entries were 'spelled' by the time Baudouin de Courtenay explored the concept (Baudouin de Courtenay, 1972). Baudouin believed that phonemes were *intentions*, but that speakers often (perhaps normally) 'missed' their intentions due to *divergences*, some of which were automatic and unconscious (psychophonetic), and others of which were (we would now say) conventionalized (paleophonetic). Kramský (1972, 27) describes Baudouin as being interested in "the representational area of individuals' linguistic consciousness—the psychic equivalent of a sound" (Baudouin, 1972 [1894]: 152) (what we would now call an abstract mental image).

A countervailing conception, originally espoused by Saussure, argued that phonemes, like all other units in language, were defined by 'otherness', as existing solely in terms of their *value* in the system, as he defined for the future of many fields the notion of *structuralism* (Saussure, 1974 [1916]). Furthering this view was the view developed by the Prague School, as exemplified by Trubetzkoy (1939 [1969]), who introduced the notion that phonemes were defined by *features* of otherness, and that a phoneme was made up of a list of those features. Bloomfield, a proponent of much more concrete theorizing, nonetheless gave phonemes the same definition—a "bundle of distinctive features" (Bloomfield, 1933: 77).

Prague School linguists such as Trubetzkoy and Jakobson specifically attacked Baudouin's view of the phoneme because the Prague school believed that speculation about internal representations was inappropriate for linguists. Trubetzkoy argued for what we would now term an 'autonomous' definition of the phoneme, a structural one, because, under an appropriate division of labor, psychologists' job was to think about storage, but linguists' job was to understand systems in terms of the internally contrasting elements – "reference to psychology must be avoided in defining the phoneme since the latter is a linguistic and not a psychological concept" (Trubetzkoy 1939 [1969]: 38). Similarly, Jakobson (quoted in Kramský, 1972: 27) said Baudouin's theory called for "the disadvantageous transfer of phonological problems from the

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

firm ground of linguistic analysis to the hazy area of introspection and their being made on such unknowns as the psychic impulses of the speaker...".

A counterview, during the same period of the development of twentieth-century linguistic theory was that of Jones (Jones, 1967), who argued that Jones argued that any given phoneme is a 'family' of related sounds.

Edward Sapir, in his classic article on the subject (Sapir, 1972) made a strong case for a similar view to that of Baudouin, that phonemes were actual *mental* images, but that the production of the images diverged from their targets due to 'the application of absolutely mechanical phonetic laws' (p. 25). As evidence he cited such observations as the fact that his Southern Paiute consultant, whom he had taught to write Papago, insisted on writing what Sapir clearly heard as voiced consonants as voiceless, because the voicing was allophonic. Notice, incidentally, that defining phonemes through the tools of minimal pairs, and the concentration on 'otherness' does not occur within Sapir's conception of the phoneme. While minimal pairs may have been useful for the linguist attempting to break into a system s/he does not speak, Sapir did not appear to believe that native speakers use that methodology to learn the language in the first place.

Within the American Structuralist tradition the more 'rigorous' Bloomfieldians made no commitments to the psychologically real (as opposed to analytically useful) nature of phonemes, and Twaddell is famous for arguing that any attempt to speculate about the contents of the mind was akin to kindling 'a fire in a wooden stove' (Twaddell, 1935: 9) but Kenneth Pike, from a competing group of structuralists (and one of Sapir's students) argued against the classic Trager-Smith phonemicization of American English in part because he found it very difficult to get linguistics students to understand the vowel system Bloch/Trager/Smith proposed (Pike, 1947).

The modern 'generative' history of the phoneme is well-known, from Halle's classic attack (Halle, 1959), as well as the basic statement of generative phonology (Chomsky & Halle, 1968). Essentially, generative phonology continued the Saussurean/Trubetzkoyan contrastive definition.[2] There was a time during the late twentieth century development of phonemic theory (not what it was called, of course, but what it was, nevertheless) when the notion of *distinctive features* was pushed as the definition of the phoneme to an extreme logical conclusion. It was

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

suggested that there could be phonemes (one per language) which were essentially the 'default' phoneme for the language. The default phoneme was defined by no features at all (one version of this was *Underspecification Theory* (see, for example Archangeli, 1988 and Steriade, 1996).

Within contemporary generative phonology the actual psychological status of the phoneme is not settled. There are those working within Optimality Theory who are simply not worried about the psychological reality of phonemes (or of the model in general),[3] and others who attempt to ground all of their theoretical constructs in physiological or psychological concepts (see e.g. the readings in Hayes, Kirchner & Steriade, 2004)

A non-generative stream that was very influential in the seventies and eighties, but has since been largely eclipsed, was Natural Phonology (Donegan & Stampe, 1979; Stampe, 1987). Natural Phonology argued that phonemes were concrete mental images of sounds, but that, as Sapir and Baudouin had argued, actual speech diverged from these stored sounds, and that the process of speaking included the real-time calculation of the divergences (and, in fact, that the divergences were dependent on speech situation, degree of precision intended by the speaker and the frequency of the utterance).

More significant, Stampe argued that speech *perception* was mediated by the same process, as hearers (who were also speakers) went through a process of 'sympathetic reconstruction' of what they heard, going through a process that could be paraphrased something like this: 's/he just said [x], which sounds like what would have come out if I had intended to say /y/'. Such a model explains many well-known phenomena across a large number of fields, including, for example, the facts that second language acquisition theorists had subsumed under the rubric of 'interference', as well as the 'phoneme restoration' facts (Samuel, 1996), in which subjects report hearing sounds that have been 'surgically' removed from a waveform, even if they are listening fairly carefully. The model is also not crucially dependent on the discovery of minimal pairs, as phonemes are units of intended perception, not classes of existing sounds.

The notion of the perception of intentions, incidentally, is not foreign to psychologists—although Stampe and Donegan did not explicitly refer to his work, it is clear that their view is similar to the Gibsonian school of perception (Gibson, 1979), which argued that organisms are attuned to the nature of their environment such that perception is of higher-level entities, and that

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

the perceptual system is designed to automatically reconstruct the objects 'behind' the percepts that impinge on the sense organs. Thus, if appropriately configured sets of moving lights are projected in a dark room people (and spiders, as Gibson showed in an elegant experiment) actually perceive moving objects, not coordinated moving sets of lights. The fact that little glowing lights attached to a person moving around in a totally dark room is perceived as a person (another experiment developed by Gibson) further emphasizes the importance of interpretation of percepts as unified entities. Similarly, the Haskins school of speech perception argued (see, for example Liberman & Mattingly, 1985) that people can be argued to perceive vocal tract movements, not abstract sound patterns.

II. THE PHONEME IN COGNITIVE GRAMMAR

The Cognitive Grammar revolution, which began in 1986 with the twin works of Lakoff (1987) and Langacker (1987) changed the way (at least some) linguists looked at such fundamental linguistic concepts as rules, lists and representations. It argued strongly for an insistence on psychological grounding for all linguistic units, claiming that there are no linguistically-specific principles at all. Lakoff and also Johnson (1987) argued that linguistic knowledge represented *embodied* experience, which Nathan (1999) argued meant that phonemes were represented as articulatory and acoustic mental images. Furthermore, Langacker argued, from the beginning, that all linguistic theoretical entities must be based on properties of the linguistic data without reliance on a priori linguistic categories such as abstract linguistic features that generative grammar has been forced to assume are innate categories.

Nathan (1986, 1996, 1999) argued that the Stampean Natural Phonology view of the nature of phonemes (and the ontological status of natural processes) was compatible with the Cognitive Grammar commitments to explanation and non-autonomy of theoretical constructs, arguing that phonemes were basic-level prototype structure categories (in the Roschian sense—(Rosch 1975, 1978)) and that natural processes were image-schema transformations analogous to those explored in some depth in Lakoff (1987). A similar claim, although not made within the Cognitive Grammar framework, had been made by Jaeger (Jaeger 1980a, b)

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Other Cognitive Grammar theorists have taken issue with the fundamental notion of representation of abstract sound images, arguing that such abstract categories were inconsistent with the fundamental insistence of Cognitive Grammar on concrete, experience-based grounding for all linguistic units. The development of what has come to be called 'usage-based' linguistic theories (Langacker, 1988, Bybee, 1999, Barlow & Kemmer, 2000, Bybee, 2001, Kristiansen, 2006) has argued that speakers construct phoneme categories on the basis of numerous instances of actual speech events stored, in some cases, 'on top of each other', so that commonalities arise naturally out of similarities in acoustic images, or articulatory events.

Others, working in frameworks separate from those of Cognitive Grammar, but compatible with it, such as Pierrehumbert (2002), have argued that prototype theory is not an appropriate model for storage of linguistic (or other) kinds of experience, proposing instead versions of *exemplar* theory, which, similarly to usage-based models argues that speakers store virtually all instances of everything they have heard, extracting commonalities from instances that are categorized as, in some sense, 'the same', with those commonalities often corresponding to units roughly the size of phonemes, but the phonemes themselves being secondary to the individual instances of individual lexical items (and, of course, larger, and perhaps smaller units). Recent work by those working explicitly within the framework of Cognitive Grammar, such as Bybee (2000, 2001), have argued that phonological theory has erred in assuming the existence of abstract phonological categories at the level of the phoneme.

Bybee has also argued that words are actually stored as individual instances, and that speakers evolve generalizations from similarities among pieces of the words, but without ever recoding the existing words in all their phonetic detail. That is, according to Bybee, phonemes are generalizations built upon existing stored entities, but do not, in any sense, change the way that words are stored. Bybee goes so far as to suggest that, for example, the allophones of phonemes in syllable-onset position may not be stored as in any sense 'the same' as those for what are traditionally thought of as the same phoneme in syllable-final position, suggesting, for example, that clear and dark [l] in English might not be categorized the same (Bybee, 2001: 88).

Although I will argue against this general view below, I simply point out here that not only do children not appear to have any difficulty spelling *leap* and *peel* with the same

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

consonants in different orders, but that it is extraordinarily difficult to get naive phonetics students even to *hear* the difference between the two sounds. Special training (such as a phonetics course) is required to hear this difference, which makes it quite different from the contrast, say between /s/ and /z/. Although these two sounds are frequently written the same (*loose* vs. *lose*, for example), children have no trouble learning the letter <z>, or using to represent /z/, and notice the jocular spelling of the plural with a <z> in, for example, the illegal file-sharing world of computer enthusiasts (illegal files are referred to as 'warez', Urbandictionary.com 'warez').

It is certainly not the case that the primary school teachers and parents who teach their children to spell are aware of these differences, but if the differences are as great as any other lexical difference we would predict clear and dark [l] would be easy to hear, not to mention other differences, such as aspiration. Children sometimes incorrectly categorize voiceless stops after /s/ as voiced (although very seldom, because that would violate very natural phonotactic restrictions which are exceptionless in English), but it is never difficult to teach children to spell 'stop' with a <t> (see Treiman, 1985, 1993 for some discussion), but, on the other hand, it is *very* difficult to get phonetics students to hear final devoicing in American English, so that they can hear that *bread is* actually pronounced with a voiceless unaspirated stop by most speakers most of the time.

III. WHY PHONEMES ARE OPERATIONAL MENTAL CATEGORIES

III.1. Phonemes and writing systems

As a counter to the view that phonemes, especially defined as recurring identical (even though they are not) sound entities, are an invention of linguists, it should be pointed out that the notion of a small number of segment-sized recurring units is one of the oldest ongoing concepts in human culture. Extensive discussion of writing systems can be found in Daniels and Bright (1996).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Recall that the Phoenecians began writing symbols that stood for consonants approximately 1800 BCE (there is some evidence that they got the practice from an earlier Egyptian variant on hieroglyphics). Their innovation was to use pre-existing symbols to represent the onset of the first (normally stressed) syllable of the word, so, for example, the symbol for 'ox' /ʔalep/ represented glottal stop, and the symbol for 'water' /mem/ represented /m/ (and the modern Roman uppercase letter has not changed much from that early symbol, amusingly enough). The alphabet (technically an 'abjad', because there were only symbols for consonants when first adopted) spread wildly around the Middle East in the second millennium BCE, becoming the basis for Hebrew, Arabic, Greek and finally Latin alphabets. Since that time, of course, it has spread throughout the world, becoming the basis for the majority of writing systems currently in use.

Of those not using Phoenician-based alphabets Korean uses a phoneme-based alphabet,[4] Japanese uses syllabaries (in addition to the Chinese-based logographic system), and, of course, the Chinese logographic system is still widely used by all Chinese speakers as well as still being used by older people in Korea, Vietnam, and, as mentioned above, is still an essential part of the overall writing system in Japan. All other living writing systems are variants of either alphabets or syllabaries (Mongolian, Coptic, Cree/Inuit). Thus, the most popular kind of writing systems are based on the assumption that there is a small number of repeatable, meaningless sound units, although opinions differ as to the exact size. Languages with complex syllable structure (essentially, any language with codas of any kind) normally choose an alphabet rather than a syllabary.[5]

The fact that, in culture after culture, language after language the writing system that survives is alphabetic (or, occasionally, syllabic) tells us that the psychological reality of understanding speech as made up of segment-sized, meaningless and recombinable units is very strong. It is true that the acquisition of literacy is a non-trivial task, but the fact that the vast majority of young children across many cultures learn to write an alphabet within less than a year suggests that the phoneme, a linguistic concept, has considerable psychological validity.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

III.2. Phonemes and speech errors

There are a number of facts about how people process language sounds that any phonological theory will need to account for. Most of these are well known, but it seems useful to rehearse them here, as it will be necessary for any model based simply on generalizations of individual usage events to account for their occurrence as well as the more traditional representational models (especially the phonetically realistic ones, such as Natural Phonology).

One area that requires serious attention is the existence of phonological processing errors that are incomprehensible unless we can assume that speakers are dealing, online, in the process of speaking, with real units at the level of abstraction the analyses by Sapir/Baudouin/Stampe suggest. Consider two examples the author gathered while participating in recent conversations, one in a meeting about computer programs, the other in a personal conversation.

The first example illustrates that syllable units such as the rhyme participate in ongoing speech production. A native speaker of American English, while aiming at

1 [bæ̃nɚ pʰɑɚʔnɚ] 'Banner partner'

produced the following:

2 ['bɑɚʔnɚ ˌpʰæn…]

To explain this we note the syllable boundaries. Syllabification of the two words is as follows:

3 /b**æn**.ɚ/ and /p**ɑɚt**.nɚ/

What appears to have 'switched places' are the corresponding *rhymes* of the stressed syllables, leaving the onsets in place. Unless rhymes are actual linguistic units involved in the planning and production of speech we cannot explain why we find 2.

It goes without saying, of course, that rhymes are also crucial in the construction of poetry and song, and that both activities do not require literacy. While twentieth century song writers no doubt are generally able to read and write, there are many non-literate cultures (both contemporary and historical) in which poetry and song based on rhyme schemes, or on alliteration—the identity of *onsets*—is a standard part of the culture. And it is also certainly the case that preliterate children have no trouble appreciating nursery rhymes long before they can spell.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

A more complex example, captured in the week this paper was being written is the following:

4 Zeno's Paradox -->['zi,naksəz...]

Here what has exchanged is again the rhyme, but in this case, the rhyme of the secondarily stressed second syllables (or perhaps feet). That will account for the output [zinaks]. However, we also find evidence for the online, real time selection of the appropriate allomorph, [əz], rather than what would have come out without the speech error, namely [z]. This is exactly the same kind of "absolutely mechanical phonetic laws" that (Sapir, 1972: 25) referred to.

A final example, discussed in earlier publications, was the error produced when a speaker aimed at 'whatever was...'. Surely such a phrase, containing the word 'whatever' is as good a candidate for a stored lexical unit as anything one could imagine. Note that the word 'whatever' is always said, in American English, with a tap. However, the actual utterance was:

5 [waʔwəzɛvəɾ]

We cannot understand this utterance without admitting that the flap in 'whatever' is not stored as such. It cannot be, else where would the glottal stop have arisen? The glottal stop is, of course, the appropriate instantiation of the /t/ phoneme in 'what',[6] when it occurs preconsonantally.

Current models such as those proposed by Bybee and others, in which there is little online construction of speech from constituent units, cannot account for speech that appears to have been constructed from more abstract units in real time.

III.3. Evidence from vowel shifts

Another well-known fact about phonological behavior that requires some explanation comes from historical change. The widely studied sound change known as the Northern Cities Vowel Shift involves a chain of changes (whether it is a push or pull chain is not important for our purposes). The change, which applies in the American cities surrounding the Great Lakes (Buffalo, Cleveland, Detroit, Chicago, and, to some extent, St. Louis) applies to the lax vowels, and has the following character:

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

5 ɔ > ɑ > a, æ > ɛə, ɛ > ʌ

Labov (1996) describes this change as follows:

> The shift begins when /æ/, the vowel of *cad*, moves to the position of the vowel of idea /iə/ (1). The vowel /o/[7] in *cod* then shifts forward so that it sounds like cad to speakers of other dialects (2). /oh/ in *cawed* moves down to the position formerly occupied by *cod* (3), /e/ in *Ked* moves down and back to sound like the vowel of cud (4), *cud* moves back to the position formerly occupied by *cawed* (5), and /i/ in *kid* moves back in parallel to the movement of /e/ (in section entitled 'Chain Shifts'—no page numbers provided)

This change applies to every word containing the relevant sound, although other sound changes appear to work their way through the lexicon via one route or another. What is important about such sound changes is that they apply to every sound, and the same sound does the same thing in every word (with the exception, for lexical diffusion, that some words simply 'do nothing' and do not participate in the change). Incidentally, all sound changes (see e.g. Labov, in press) represent the changes that establish classic dialect and language families and cause linguists to set up genealogical trees).

How can this be accounted for in theories in which phonemes are abstractions generated from repeated individual instances of words containing the sound? If, for example, /æ/ is an abstraction discoverable by, in some sense, examining every instance of every word a speaker has ever heard that contains that sound, how can we explain the fact that the same speaker (always a child in the case of true 'transmission', not 'diffusion') changes every single word. If the categorization of the sounds is secondary to their storage we need to assume that each word is somehow 'indexed', and that speakers need to do a 'find and change' (analogous to what we do in a word processor). If, on the other hand, phonemes are the actual units in which words are 'spelled', perhaps as some kind of unification of a set of articulatory instructions and a set of (rough) formant patterns, then a single change in the phoneme itself, however it is represented, will lead to its being activated everywhere, since there is only one sound to change.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

A second corollary to this view (namely that when children learn words they actually store them as sequences of this small set of basic units (as Taylor, 2006 also argues), is that we can understand how chain shifts can occur. Note that the Northern Cities Vowel Shift is such a chain shift. What this means is that in some sense the /ɑ/ vowel is interacting with the /æ/ vowel, which is interacting with the /ɛ/ vowel (I use the neutral term 'interacting' to avoid a commitment to 'push' versus 'drag' chains, a point that is unimportant to my argument).

If words are simply stored as individual items, perhaps with some cross-classifying 'index', then it would make no sense to say that the /æ/ vowel is 'pushing on' the /ɛ/ vowel, since there *is* no category of /æ/ vowels to do any pushing. It is certainly meaningless to say that the word 'bag' is pushing on the word 'ready'. We need to have some sense of active cognitive processing, involving categorization and production in real time, not 'recollected in tranquility' if we are to understand how phonemes could move in a coordinated manner, as we know they have done numerous times in the past, and continue to do as we study ongoing living speech. It is only a within a model in which phonemes have some independent, real existence, but are also instantiated in each word that contains that we can understand how real regular sound change, either historical or ongoing could take place.[8] That is, we need the notion of an inventory of actual phonemes and a lexicon of words 'spelled' in those phonemes.

III.4. Evidence from speech processing studies

Yet another way in which phonemes appear to be real, mentally coherent categories in which speakers operate in real time as they speak and hear their language comes from the fact, discussed by Cutler (2002), that allophonic processes normally do not show lexical conditioning:

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

> Studies of coarticulation reveal regularities which are determined by phonemic environment—the gestures which correspond to /k/ are different if the following vowel is high front /i/ rather than low back /ɔ/, for instance. Such studies have not revealed a role for the word itself as a determiner of regularity—high frequency words such as *key* and *cause* and low frequency words such as *kiwi* and *caucus* show the same patterns of variation. Without some expansion of the episodic modeling framework beyond word-specific phonetics, such regularities must presumably be ascribed to chance (p. 287).

A recent study of the perception of stress in contrasting language types emphasizes the importance of recognizing that some facts about words are simply not stored, no matter how many times the words may be repeated, and no matter how many 'obvious' the fact is to an outside observer. It is well known that French stress is completely automatic, falling on the final full vowel (needed to exempt schwas) of a breath group (Schane, 1968b). As such, French speakers regularly produce stressed syllables, but apparently do not store stress at all.

In order to test this assumption, Peperkamp and Dupoux (2002) attempted to teach French speakers (as well as Finnish, Hungarian, Polish and Spanish speakers) to repeat non-words that contrasted in stress placement alone. They found that French and Finnish speakers made significantly more errors in stress placement than did Hungarian speakers, while Polish and Spanish speakers did best. These languages can be rated along a scale of stress predictability—Hungarian, although having initial stress, has exceptional borrowed words, while Polish and Spanish have a stress 'window' of the well-known kind (somewhere in the last three syllables, depending on phonology, morphology and lexical effects). Speakers of French and Finnish simply can not get the stress right at all—they are, in the words of the authors, 'stress deaf'. But surely speakers of these languages *hear* stress all the time, and, if usage-based theories are correct, must be storing it. Nonetheless, they seem incapable of hearing it in the 'wrong' place, where their language forbids it, even if the task is simply to repeat a word heard immediately before.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Taylor (2006) suggests that phonemes are basic-level categories (following a discussion in Nathan 1986 and elsewhere). Why speakers should feel the need to store speech in terms of a small number of production/perception units remains a question that many continue to ask, and many now contest. I think a suggestion for the reason might be found in the original notion of basic level category, as proposed by Rosch, and elegantly summarized in a recent work on how basic musical motifs (such as the opening phrase of Beethoven's *Fifth Symphony*) constitute basic level categories. Zbikowski (2002) suggests:

> Rosch and her associates suggested that two contrasting principles influence the taxonomic level at which people prefer to categorize. The first is the efficiency principle, according to which people prefer to minimize the number of categories they must consider in making a categorization....The second principle is the informativeness principle, according to which people tend to maximize the informativeness of their categorizations. Since the most information about any entity is found at the most specific level of a taxonomy, you would use grand piano to categorize the thing sitting in the living room. Rosch and her associates proposed that the intermediate level of a taxonomy (in this case piano) optimizes both efficiency and informativeness and is thus the preferred level for basic categorization. A number of empirical operations converge at the basic level. The basic level is the highest level whose members have similar and recognizable shapes; it is also the most abstract level for which a single mental image can be formed for the category. The basic level is also the highest level at which a person uses similar motor actions for interacting with category members. The basic level is psychologically basic: it is the level at which subjects are fastest at identifying category members, the level with the most commonly used labels for category members, the first level named and understood by children, the first level to enter the lexicon of a language, and the level with the shortest primary lexemes' (pp. 32-33).

Similarly, Mompeán (2006) shows that categories at the level of abstraction roughly equivalent to the phonemic level are most easily perceived and stored.

III.5. Evidence on the recoding of stored examples

It is indeed the case that our ability to store things we have heard is much greater than those who worry about the 'poverty of the stimulus' would credit. Pierrehumbert (2002) cites evidence that

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

speakers recognize new words more easily the next time they hear them if they hear them in the voices of the speakers who introduced them to those words. She argues on that basis that phonemes are simply generalizations over numerous stored instances, without the need for much in the way of recoding (although she allows for the possibility that some repeated instances are similar enough to others that the new instances may not be stored separately.

However, there is evidence against this view, and in favor of a view in which learners store speech 'as they would say it', rather than behaving as Taylor says, 'like tape recorders' (Taylor, 2006: 45). It is certainly the case that we can hear, and remember how a specific speaker sounds while saying a particular word, in some detail. Research on the perception and storage of music also argues that we have stored information about the absolute pitch of songs we have learned (say by hearing them repeatedly on the radio). Levitin (2006: 149) found that people who are asked to sing songs they have learned in that way normally start 'at or very near the absolute pitches of their chosen songs'. This would appear to argue that children (and presumably adults) learn new words in exactly the same way. But when children learn to speak there is no evidence that they pronounce words they learn from their fathers on a lower pitch than the words they learn from their mothers, and, of course, in general they pronounce all words at a higher pitch than either (in part, of course, because their vocal cords are smaller). But a moment's introspection will tell us also that when we are introduced to someone who has an unusual name, we do not repeat it on a different pitch if the person is female than if they are male. Clearly some instantaneous recoding has gone on. And, of course, the same recoding is quite possible. Although Levitin did not deal with this issue, it is very unlikely that, if someone with a bass voice is asked to sing a song he learned initially from a soprano that he will attempt to match the soprano range. Instantaneous transposition is something anyone can do (I am speaking here, of course, of actual singing—*writing transpositions*, in the sense of musical notation is a difficult

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

skill that takes several years of formal musical training). And, of course, we instantaneously repeat someone's name in our own dialect when we are introduced: when someone with the Northern Cities Dialect shift introduces themselves to me as [mɛət] I reply 'glad to meet you [mæt]'. Not only would it be rude to imitate the Detroit vowel, I simply cannot do so with the authentic vowel required, and my students find it amusing to hear me try.

III.6. Evidence from child language acquisition

Although much has been written on this topic it seems important to mention once again that there are numerous reports of children systematically substituting one phoneme for another in every word containing that phoneme. Smith (1973) is a classic study, for example, but a bilingual child that my colleagues and I have been studying systematically replaced all mid vowels (/e/, /o/) in his Spanish with high vowels for several months. Similarly, a child of my acquaintance currently learning English regularly replaces words containing voiceless dental fricatives (/θ/) with labiodentals (/f/). This kind of replacement, virtually universally reported, poses two problems for any usage-based theory. First, where do children learn to replace what are almost always *more marked* segments with *less marked* segments. They certainly are not learning it from their surroundings, because these are not only non-standard replacements, they are non-existent in the ambient language. And secondly, unless children are storing the words they attempt *with the phonemes as separate representations*, how else can they know which words contain the relevant sounds. If we assume that phonemes are, in some sense, 'calls to motor routines', rather than simply linguists' classification schemes, we can explain why children alter the routines, leading to replacements in every relevant word.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

IV. CONCLUDING THOUGHTS

How can we sum up where the phoneme stands early in the twenty-first century among functionalist and other not strictly structuralist phonologists? It is certainly reeling under the combined assault of a large group of Cognitive Grammarians (Bybee, Kemmer, Langacker and others) and a smaller set of phonetically-oriented experimental phonologists (Pierrehumbert, Beckman (2000), Port (Port & Leary, 2005) and others). They provide evidence that speakers can hear fine phonetic detail, and appear to store it long-term. On the other hand, speakers have no metalinguistic access to that information. Unless the individual (or even dialectal) details cross phoneme boundaries (specifically the phoneme boundaries of the speaker/hearer), speakers have no way to talk about the differences. Without extensive training in phonetics speakers are unable to *describe* the clear [l]'s of Irish English or the distinctive way George W. Bush pronounces 'decider'.[9] Furthermore, no language (other than the specialized tool of the IPA) provides a way of recording such details, leading one to assume that speakers feel no need to identify it. Furthermore, people have been learning to 'spell' for almost four thousand years, and, however complex that task may be, it is a task that requires no scientific equipment, and can apparently be learned 'spontaneously' at least by some children (Read, 1986).

I believe that there remains reason to believe that people *do* hear a small number of distinct sound units *in* the words that they acquire as children. Furthermore, their storage of those sound units appears to be active not only in perception but also in production not only of speech they have heard before, but also of novel utterances, and even of the non-words produced during whatever processing errors are reflected in spoonerisms and other speech errors. There are very few psychological theories that remain quite as robust four thousand or so years after they were first invented, but So, although many are aiming at it, this four thousand year old target—the perception and production of speech sounds by naïve speakers in terms of recurring, identical

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

segments, which differ apparently without notice in different contexts, in both language-specific and universal ways—in short, the phoneme, appears to be resisting calls for its demise.

NOTES

1. I would like to thank Margaret Winters and José A. Mompeán for their comments on earlier drafts of this paper.

2. It should be pointed out, however, that even in the midst of the more orthodox generative view there were generative linguists, such as Schane (1968a, 1971) and Kiparsky (1982) who 'reverted' to a more traditional 'autonomous' view of the phoneme, not so much in the sense of autonomy from mental reality (as Trubetzkoy and Jakobson preferred) but in the Chomskyan polemic sense of not necessarily related to morphophonemic alternations, and derivable solely by distributional facts. It is important to note, however, that Schane and Kiparsky were arguing for the *psychological reality* of the 'autonomous' phoneme.

3. Consider, for example, the fact that the basic reference to OT, Prince and Smolensky (2004 [1994]) does not discuss the issue at all. Much of the literature on learnability of OT, in addition, is based on theoretical discussions of learnability as algorithms, rather than dealing with actual language input.

4. While the symbols in Korean are arranged into syllabic units, each component symbol represents a (somewhat abstract) phoneme. See, for example, King (1996).

5. Hittite attempted to use a syllabary, with mixed results. There was great inconsistency on how CVC structures were written—sometimes CV-V-VC, other times CV-VC, and yet others CV-V-V-VC. See Gragg (1996).

6. This is a fact about American English, but the speaker (the author) is a native speaker of that dialect for whom /t/ is always pronounced as glottal stop preconsonantally.

7. Labov uses the Trager-Smith transcription to describe American English. Thus he uses /o/ to represent what was historically /a/, and /oh/ to represent what is, in the dialects that have this sound, approximately /ɔ/

8. Not only does Labov assume in the articles mentioned above that there are real, categorically regular sound changes, but he is convinced that they exist, and that they serve as a counterexample to at least some part of the totally usage-based models that Cognitive Grammar has recently taken to heart (Labov, p.c.).

9. Note that the stigmatized pronunciation [nukjələ˞] is *phonemically* very distinct from the standard pronunciation [nuklijə˞], making comment on Bush's habitual pronunciation fodder for his critics.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

REFERENCES

Anderson, S. R. (1985). *Phonology in the twentieth century: Theories of rules and representations*. Chicago: University of Chicago Press.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, 5(2), 183–208.

Barlow, M., & Kemmer, S. (Eds.) (2000). *Usage-based models of language.* Cambridge: CUP.

Baudouin de Courtenay, J. (1972). An attempt at a theory of phonetic alternations. In Edward Stankiewicz (Ed. & Trans.), *A Baudouin de Courtenay anthology: The beginnings of structural linguistics*. Bloomington/London: Indiana University Press, pp. 144-213.

Beckman, M. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development*, 71(1), 240-249.

Bloomfield, L. (1933). *Language.* New York: Holt.

Bybee, J. L. (1999). Usage-based phonology. In M. Darnell, E.-A. *et alia* (Eds.), *Functionalism and formalism in linguistics, I: General papers; II: Case studies.* Amsterdam: Benjamins, pp. 211-242.

Bybee, J. L. (2000). The phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of Language*. Stanford: Center for the Study of Language and Information, pp. 65-86.

Bybee, J. L. (2001). *Phonology and language use.* Cambridge & New York: CUP.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper and Row.

Cutler, A. (2002). Phonological processing: Comments on Pierrehumbert, Moates et al., Kubozono, Peperkamp & Dupoux and Bradlow. In C. Gussenhoven & N. Warner (Eds), *Laboratory Phonology 7*. The Hague: Mouton de Gruyter, pp. 275-296.

Daniels, P. T. & Bright, W. (Eds.). (1996). *The world's writing systems*. New York: OU.

Donegan, P. J., & Stampe, D. (1979). The study of Natural Phonology. In D. Dinnsen (Ed.), *Current approaches to phonological theory.* Bloomington: Indiana University Press, pp. 126-173.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Gragg, G. B. (1996). Other languages. In Daniels, P. T. & Bright, W. (Eds.). (1996). *The world's writing systems*. New York: Oxford University Press, pp. 65-70.

Halle, M. (1959). *The sound pattern of Russian.* The Hague: Mouton.

Hayes, B., Kirchner, R. M., & Steriade, D. (Eds.). (2004). *Phonetically-based phonology.* New York: CUP.

Jaeger, J. J. (1980a). *Categorization in phonology: An experimental approach*. Ph.D. dissertation. University of California, Berkeley.

Jaeger, J. J. (1980b). Testing the psychological reality of phonemes. *Language and Speech*, 23, 233-253.

Johnson, M. (1987). *The body in the mind*. Chicago: University of Chicago Press.

Jones, D. (1967). *The phoneme: Its nature and use.* Cambridge: Heffer.

King, R. (1996). Korean Writing. In *The World's Writing Systems*. In P. T. Daniels & W. Bright (Eds.), Oxford: Oxford University Press, pp. 218-227.

Kiparsky, P. (1982) How abstract is phonology?. In Paul Kiparsky (Ed.), *Explanation in Phonology*. Dordrecht: Foris, pp. 119-164

Krámský, J. (1974). *The phoneme: Introduction to the history and theories of a concept*. München: W. Fink.

Kristiansen, G. (2006). Towards a usage-based phonology. *International Journal of English Studies*, 6(2), 107-140.

Labov, W. (1996). The organization of dialect diversity in North America http://www.ling.upenn.edu/phono_atlas/ICSLP4.html

Labov. W. (2007). Transmission and diffusion. To appear in *Language*, June 2007

Lakoff, G. (1987). *Women, fire and dangerous things. What categories reveal about the mind.* Chicago: University of Chicago Press.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar.* Stanford: Stanford University Press.

Langacker, R. W. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics*. Amsterdam: Benjamins, pp. 127–161.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Levitin, D. J. (2006). *This Is your brain on music: The science of a human obsession*. New York: Dutton

Liberman, A., & Mattingly, I. (1985). The motor theory of speech perception revisited. *Cognition*, 21, 1–36.

Mompeán, J. A. (2006). The phoneme as a basic-level category: Experimental evidence from English. *International Journal of English Studies*, 6(2), 141-172.

Nathan, G. S. (1986). Phonemes as mental categories. In *Proceedings of the 12th Annual Meeting of the Berkeley Linguistics Society*, 12, 212–224.

Nathan, G. S. (1996). Towards a Cognitive Phonology. In B. Hurch & R. Rhodes (Eds), *Natural Phonology: The state of the art*. Berlin: Mouton/de Gruyter, pp. 107–120.

Nathan, G. S. (1999). What functionalists can learn from formalists in phonology. In *Proceedings of the Symposium on formalism and functionalism.* Amsterdam: Benjamins, pp. 305–327.

Peperkamp, S., & Dupoux, E. (2002) A typological study of stress 'deafness'. In C. Gussenhoven & N. Warner (Eds) *Laboratory Phonology 7*. The Hague: Mouton, pp. 203-236.

Pierrehumbert, J. (2002). Word specific phonetics. In C. Gussenhoven & N. Warner (Eds), *Laboratory Phonology 7*. The Hague: Mouton de Gruyter, pp. 101–139.

Pike, K. (1947). On the phonemic status of English diphthongs. In V. B. Makkai (Ed.), *Phonological theory: Evolution and current practice*. New York: Holt Rinehart and Winston, pp. 145–151.

Port, R.F. & Leary, A. P. (2005). Against formal phonology, *Language*, 81(4), 927-964.

Prince, A. & Smolensky, P. (2004 [1994]). *Optimality Theory: Constraint interaction in Generative Grammar*. Oxford, UK: Blackwell. Originally published as Technical Report Number 2 of the Rutgers Center for Cognitive Science.

Read, C. (1986). *Children's creative spelling*. London, England: Routledge & Kegan Paul.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–233.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B.B. Lloyd. (Eds.) *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum, pp. 27–48.

Samuel, A. (1996). Phoneme restoration. *Language and Cognitive Processes*, 11, 647–653.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Sapir, E. (1972). The psychological reality of phonemes [La réalité pschologique des phonèmes] (Translated and reprinted from *Journal de psychologie normale et pathologique* (1933), 30, 247–265). In V. B. Makkai (Ed.), *Phonological theory: Evolution and current practice.* New York: Holt Rinehart and Winston, pp. 22-31

Saussure, F. d. (1974). *Course de linguistique générale. Edition critique préparée par Tullio de Mauro.* Paris: Payot.

Schane, S. A. (1968a). On the non-uniqueness of phonological representations. *Language*, 44, 709-716.

Schane, S. A. (1968b). *French phonology and morphology*. Cambridge, MA: MIT Press.

Schane, S. A. (1971). The phoneme revisited. *Language*, 47, 503-521.

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge: CUP.

Stampe, D. (1987). On phonological representation. In W. U. Dressler *et alia* (Eds), *Phonologica 1984*. London: Cambridge University Press, pp. 287–300.

Steriade, D. (1996). Underspecification and markedness. In J. A. Goldsmith (Ed.), *The handbook of phonological theory.* Cambridge, MA: Blackwell, pp. 114-174.

Taylor, J. R. (2006). Where do phonemes come from? A view from the bottom. *International Journal of English Studies*, 6(2), 19-54.

Treiman, R. (1985). Spelling of stop consonants after /s/ by children and adults. *Applied Psycholinguistics*, 6, 261-282.

Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York: Oxford University Press.

Trubetzkoy, N. S. (1939). *Gründzüge der phonologie [Principles of Phonology]* (C. Baltaxe, Trans. and Ed.). Prague [Los Angeles]: *Travaux du cercle linguistique de Prague* [University of California Press].

Twaddell, W. F. (1935) *On defining the phoneme. Language*, Vol. 11, No. 1, Language Monograph No. 16.

'warez', Urban Dictionary http://www.urbandictionary.com/define.php?term=warez

Zbikowski, L. M. (2002) Conceptualizing music: Cognitive structure, theory and analysis. Oxford: Oxford University Press.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

Välimaa-Blum, Riitta. 2005. *Cognitive Phonology in Construction Grammar: Analytic Tools for Students of English.* Berlin: Mouton de Gruyter. 271 pages. ISBN 978-3-11-018608-6

JOHN. R. TAYLOR
University of Otago

I approached this book with a good deal of anticipation. Cognitive linguistics is still very much focused on the study of conceptual structures, phonology being largely ignored or even considered not to fall within its purview. This neglect of phonology is very much to be regretted. The phonology of a language is no less amenable to a cognitive treatment than the study of word meaning, for example, in that pronunciations, just like word meanings, have to be mentally represented. Similar kinds of issues arise in the two areas, too. An important theme in lexical semantics has concerned the amount of polysemy than needs to be postulated, and how to strike a proper balance between mentally represented meanings on the one hand and context-dependent readings on the other. Analogously, phonologists need to enquire about how word pronunciations are mentally represented and how these representations relate to the pronunciation variants that occur in speech events. As the title of the present volume indicates, the notion of construction may be as relevant in phonology as it is in syntax. This is most obviously the case with regard to patterns for word formation. But even such a strictly phonological construct as the syllable is also amenable to a constructionist account, in that admissible syllables in a language need to conform to more abstractly characterized 'syllable constructions' (Taylor, 2004). A book, therefore, on cognitive phonology, especially one that promises to incorporate phonology within construction grammar, is most timely. However, I found the book under review rather disappointing, in several respects.

Let us consider, first, its contents and coverage. Leaving aside Ch. 1 (which offers a brief overview of cognitive linguistics, with some remarks on constructions and construction

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

grammars), most topics addressed in the book are basic issues that are covered in the many excellent introductions to English phonology that are already on the market. Ch. 2 discusses the consonant and vowel articulations of English, treating such matters as place and manner of articulation, vowel height, and so on. Ch. 3 accepts without question the validity of the phoneme concept, with sections on minimal pairs, complementary distribution (clear and dark 'l'), and free variation. Ch. 4 deals with morpheme alternations (*music-musician*, and the like; more on this below). Ch. 5 addresses stress location in English words. The topic is approached mainly from the point of view of various suffixes and prefixes and their effects on stress location within the derived word. Ch. 6 deals with intonation, but barely goes beyond the basics of the Hallidayan analysis (Halliday, 1970), introducing such topics as tone unit, tonic syllable, rising vs. falling tone, and their 'default' associations with statements and two kinds of questions (polar and *wh*-). It is acknowledged that polar (yes-no) questions do not always have a rising intonation as per the default, also that intonation is very much a matter of information structure and speaker-hearer relations. But these matters are not pursued, nor are they illustrated with any actual discourse examples.

Given that the topics addressed in the book are all rather basic, and to be found in just about any introductory textbook, our attention must go to the supposedly 'cognitive' perspective which the author adopts. In various places in the book, the author cites experimental evidence (e.g. Pierrehumbert, 2001; Pisoni *et al.*, 1985) to the effect that speakers store words in rich phonetic detail, not at all in an abstract, or underspecified format. Pairs such as *music* and *musician*, *divine* and *divinity*, *wife* and *wives*, would accordingly be mentally represented in a form which is schematic for the range of pronunciations which these words receive in utterance events. A speaker may, no doubt, become aware of the phonetic correspondences between these pairs, and may even notice the parallels between e.g. *wife/wives*, *leaf/leaves*, *life/lives*. This would allow the abstraction of a schema which accommodates the correspondences between (certain) singular and plural nouns; the schema, of course, takes the place of what in other theories would be regarded as a rule for plural formation. This, at least, is the 'cognitive linguistic' tack that I would take on the matter (Taylor, 2002). The author, however, proposes an amalgam of rich phonetic representation on the one hand and a three-level derivational theory on the other. There are three levels of representation. These are the morpheme level, the word level (called, curiously, the 'phonemic level'), and the utterance level (called the 'phonetic level'). The morpheme level lists the pronunciations of all of the allomorphs of a morpheme. Thus, the morpheme {wife} has the

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

dual entry [waɪf] and [waɪv], {music} has the entries [ˈmju:zɪk] and [mjuˈzɪʃ]. Representations are this level are rich in allophonic detail, yet, the author claims that they do not contain information about stress location, nor about syllabification. At the word level, morphemes are combined in accordance with various word-formation constructions. Here, words are syllabified and stress is assigned. Level 3 handles sandhi and other 'post-lexical' phenomena.

The level theory, in particular the idea that morphemes are stored in an unpronounceable form (unpronounceable, since stress and syllable boundaries are not marked), goes against the monostratal approach that cognitive linguists (and others) have pursued in the study of syntax (Langacker, 1987). The problems associated with the author's approach come to the fore in her treatment of the so-called linking 'r' in non-rhotic accents (like my own). It is assumed that orthographic 'r' is present underlyingly in the morphemic representation of words such as *far* and *farm*. The 'r', however, only surfaces at the word or utterance level if conditions are appropriate. Compare now *far (r) and near* with *ma (r) and pa*. The presence of the 'r' in the two phrases would need to be accounted for by two quite different processes, one which allows an underlying 'r' to surface, the other inserting an 'r'. Much simpler, I would have thought, and more in keeping with the facts, is to suppose that the very same process operates in each of the two cases, namely, that an 'r' may be inserted between two vowels, the first of which is non-high. (Alternatively, in this situation, a glottal stop is inserted between the vowels; see Taylor 2002). On this account, *far* would be mentally represented in the form in which it is pronounced, namely, as [fɑ:]. We would also avoid the problem of the underlying 'r' in *farm* which never actually surfaces.

Another of my reservations about the book concerns the less than perspicuous way in which generalizations about English phonology are stated and presented. Take, for example, the chapter on word stress. According to the author's theory, morphemes (and monomorphemic words) are stored in the lexicon without stress allocation, unless they happen to be exceptions to the rules. It is important, therefore, to state precisely what the rules are, in order that we know what the exceptions are. The place to begin, I would have thought, is monomorphemic words, of the kind *algebra*, *agenda*, *America*, *maintain*. The 'rules' - which could readily be represented in the form of stress-placement schemas - are nowhere clearly stated in the chapter. The rules (see Giegerich, 1992, for a good account), or 'schemas', rest on principles of syllabification and on a distinction between light and heavy syllables, which in turn presupposes the distinction between short and long vowels (alternatively, lax

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

and tense vowels). Secondly - and this is something which the author appears to have missed - nouns behave differently from words of other categories. For nouns, the Latin stress rule applies. The Latin stress rule assigns stress to the second-last syllable if it is heavy, otherwise, to the third-last syllable. The final syllable, in other words, is ignored (of course, if the noun contains only one syllable, then, perforce, it is stressed; likewise, if the noun contains only two syllables, stress must fall on the second-last, irrespective of its weight). There are, to be sure, a significant number of exceptions to the Latin rule, an important group being nouns where stress is attracted to a final syllable containing a long vowel, such as *July*, *arcade*, *magazine* (in conservative pronunciations). With respect to words which are not nouns, a variant of the Latin rule applies. A final consonant (if there is one) is ignored. Then, if the final syllable is heavy, it is stressed, if not, stress falls on the second-last syllable. Most of these topics (except, as far as I can see, the different behaviour of nouns vis-à-vis non-nouns) are, to be sure, mentioned somewhere or other in this chapter and elsewhere, but they are nowhere brought together in a succinct, and easily understood form. What we have, instead, are all kinds of seemingly ad hoc explanations put forward to account for whatever words are under consideration. Symptomatic is the fact that the author sometimes describes stress location with respect to the end of a word, e.g. as falling on the penultimate or antepenultimate syllable, sometimes as falling on the first or the second syllable, that is, with respect to the beginning of a word. This inconsistency can only add to the confusion of the beginning student.

The main thrust of the chapter, as mentioned, concerns the role of affixes. Affixes come with their own constructional schemas, which determine stress location (and also vowel weight) in the complex form. A clear case would be the adjective-forming -ic suffix, which places stress on the preceding syllable. This promising approach is compromised by the claim (p. 162) that stress in *logic* is also assigned by the same principle, even though *logic* is not a derived adjective and the final *-ic* would not normally be regarded as suffix (what is it suffixed to?). The approach is probably a consequence of the author's view that affixes do not actually have any meaning in themselves (p. 163); she therefore regards any word ending in -ic as fair game for the affixation rule. But she is then faced with the task of accounting for the 'exceptional' stress placement in *arsenic*, *rhetoric*, *turmeric*, and several more (p. 184) Actually, in terms of stress placement in non-derived nouns, these are not exceptions at all.

An approach which focuses on the role of affixation requires that complex words are properly analyzed by speakers. One of the most curious passages in the book (pp. 200-202) does address the question of morphological analyzability. What is curious about the passage is

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

that the author seems to assume that expressions are analyzable to the extent that they are entrenched; she then expresses surprise that speakers may not be aware of the internal structure of entrenched expressions. This, surely, gets it the wrong way round! As an expression gets more and more entrenched through frequency of use, it becomes automated, it acquires unit status (Langacker, 1987) and can be accessed and performed as a pre-established routine, it does not have to be assembled from its parts, nor is the contribution of the parts prominent to the user.

My final reservation about the book concerns its suitability as a pedagogical text. One of the beauties of teaching phonology is that it provides an arena, more circumscribed than syntax, where students can be introduced to techniques of 'doing linguistics'. The facts are usually evident to all, namely, how words are pronounced. But why are words pronounced as they are, e.g. with stress allocated to certain syllables and not to others, with 't'-flapping possible in some words but not in others? (because it 'feels right' to pronounce the words that way). But where does this 'feeling' come from? What are the generalizations which the data conform to? How can we test any generalizations that we might put forward? How to account, in a principled way, for pronunciation variation between speakers, and even within a single speaker? How to evaluate competing generalizations? Can we ascribe 'psychological reality' to our generalizations, and how might we decide the matter? And why should the generalizations be as they are? These are bread-and-butter issues to practicing linguists. At the very least, a book claiming to introduce students to 'analytic tools' should contain a range of study questions which challenge students to formulate, to test, and to refine generalizations, in the first instance, on 'data sets' provided by the instructor. The kinds of tasks 'for further thought' suggested at the end of each chapter of this book, which invite students to find their own examples and to do their own analysis on them, just do not work, in my experience.

REFERENCES

Giegerich, H. (1992). *English Phonology: An Introduction.* Cambridge: Cambridge University Press.

Halliday, M. A. K. (1970). *A Course in Spoken English: Intonation.* Oxford: Oxford University Press.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar, vol. 1: Theoretical Prerequisites.* Stanford: Stanford University Press.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee and P. Hopper (Eds.). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, pp. 137-157

Pisoni, D., Nussbaum, H., Luce, L., & Slowiaczek, L. (1985). Speech perception, word recognition and the structure of the lexicon. *Speech Communication*, 4, 75-95.

Taylor, J. R. (2002). *Cognitive Grammar.* Oxford: Oxford University Press.

Taylor, J. R. (2004). Why construction grammar is radical. *Annual Review of Cognitive Linguistics*, 2, 321-348.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

www.um.es/ijes

About the authors

Volume 6, Number 2, 2006

David Eddington earned his PhD from the University of Texas at Austin in 1993. he has held positions in the University of New Mexico, Middle Tennessee State University, and Mississippi State University. He joined the Linguistics & English Language Department at Brigham Young University in 2003, where he is Associate Professor and where he teaches Linguistics. His interests include experimental linguistics, phonology, morphology, and Spanish language. He is author of numerous papers in international journals like *Language*, *Brain and Language*, etc. (visit http://linguistics.byu.edu/faculty/eddingtond/profession.html for a full list). One of this most recent contributions is "Flaps and other variants of /t/ in American English: Allophonic distribution without constraints, rules, or abstractions" (Cognitive Linguistics, 2007). He is also author of *Spanish Phonology and Morphology: Experimental and Quantitative Perspectives* (Amsterdam: John Benjamins, 2004).

Helen Fraser received her PhD from the University of Edinburgh in Scotland. She has taught at the University of Edinburgh, Trinity College Dublin, and SOAS (University of London). From 1990 until the present, Helen Fraser has worked at the University of New England (Australia), where she is Senior Lecturer at the School of Languages, Cultures and Linguistics. Dr Fraser teaches at all levels in phonetics, phonology, psycholinguistics, history and philosophy of linguistics, and (Australian) English language. Her research focuses on the representation of speech sounds, pronunciation teaching, and forensic phonetics and transcription. Her publications include papers like "Phonology without tiers: Why the phonetic representation is not derived from the phonological representation (*Language Sciences*, 1997), "Constraining abstractness: Phonological representation in the light of color terms" (*Cognitive Linguistics*, 2004), or books like *The Subject of Speech Perception: An Analysis of the Philosophical Foundations of the Information-Processing Model of Cognition* (Macmillan, London, 1992).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Gitte Kristiansen received her PhD from the Universidad Complutense de Madrid. She is a Lecturer in Linguistics at the Department of English Language and Linguistics, Universidad Complutense de Madrid (Spain) since 1995, where she currently teaches courses on semantics and cognitive linguistics. Her research interests include language variation and change, cognitive sociolinguistics and cognitive phonology. She is author of "Social and Linguistic Stereotyping: A Cognitive Approach to Accents" (2001) and "How to do Things with Allophones: Linguistic stereotypes as cognitive reference points in social cognition" (2003), and co-editor of *Cognitive Sociolinguistics. Language Variation, Cultural Models, Social Systems* (Berlin/New York: Mouton de Gruyter, forthcoming) and *Cognitive Linguistics: Current Applications and Future Perspectives* (Berlin/New York: Mouton de Gruyter, 2006). Gitte Kristiansen is managing editor of the book series *Applications of Cognitive Linguistics* (Mouton de Gruyter).

Fumiko Kumashiro received her PhD from the University of California, San Diego, in 2000. Since 2000, she has worked at the University of Tokyo and Keio University. She currently teaches in the Faculty of Law and the Faculty of Letters at Keio University. Her research focuses on phonology, cognitive grammar, and corpus linguistics.

Toshiyuki Kumashiro obtained his PhD in 2000 from the University of California, San Diego. From 1993 to 1999, he was Lecturer in Japanese in the Department of East Asian Languages and Literatures at the University of California, Irvine. Since 1999, he has taught at Keio University in Tokyo, where he is now Professor at the Faculty of Law. His research interests include cognitive grammar, Japanese linguistics, grammatical relations, and English phonology. He is co-author of "Double-subject and complex-predicate constructions" (*Cognitive Linguistics*, 2003).

José A. Mompeán received his PhD in 2002 from the University of Murcia and an MA in Phonetics from University College London (UK) in 2006. Since 2004 he is Senior Lecturer in English Phonetics in the English Department at the University of Murcia (Spain), where he lectures and does research on English phonetics and phonology, cognitive approaches to phonology and pronunciation/phonetics teaching. Within the cognitive linguistics framework, he is author of papers like "Category Overlap and Neutralization: The Importance of Speakers' Classifications in Phonology" (*Cognitive Linguistics*, 2004) or "A Comparison between English and Spanish Subjects' Typicality Ratings in Phoneme Categories: A First Report" (*International Journal of English Studies*, 2001). He is currently working on a book entitled *English Phonology: An Empirical Approach*.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Geoffrey Nathan received his PhD in Linguistics from the University of Hawai`i in 1978. He worked in the Department of Linguistics at Southern Illinois University and, in 2002 he became Associate Professor in the Department of English (Linguistics Program) at Wayne State University (USA), where he is also Faculty Liaison to Wayne State's Computing and Information Technology Division. His research focuses on natural phonology and cognitive linguistics, particularly the integration of the principles of Natural Phonology with the concepts of Cognitive Grammar, and he has also published on second language phonology. He is author of numerous papers on phonology (natural and cognitive) such as "Phonemes as Mental Categories" (BLS, 1986), "On Second-Language Acquisition of Voiced Stops" (*Journal of Phonetics*, 1987), "How the Phoneme Inventory Gets Its Shape--Cognitive Grammar's View of Phonological Systems" (*Rivista di Linguistica*, 1995), "Towards a Cognitive Phonology" (In *Natural Phonology: The State of the Art*. Berlin: Mouton/de Gruyter, 1996), etc. He is also author of a forthcoming book entitled *Phonology in Cognitive Grammar*, which is currently under review.

John Taylor has held positions in Germany and in South Africa. At present he is at the University of Otago (New Zealand), and is a member of the Department of English, where he lectures on phonetics and phonology, language acquisition, and grammar. His research focuses on cognitive grammar, spatial relations and phonetics/phonology. Currently he is the managing editor of *Cognitive Linguistics Research* (Mouton de Gruyter) and is on the editorial board of the *Functions of Language* (John Benjamins), and of the journal *Cognitive Linguistics*. He is author of highly influential works in Cognitive Linguistics like *Linguistic Categorization: Prototype in Linguistic Theory* (3rd edition, Oxford University Press, 2003. 1st edition: 1989, 2nd edition: 1995) or *Cognitive Grammar* (Oxford University Press, 2002). He is also author of numerous book chapters and papers (for a complete list visit http://www.otago.ac.nz/linguistics/staff/taylor.html)

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

International Journal
of
English Studies

IJES

UNIVERSITY OF MURCIA www.um.es/engphil/ijes

Instructions to Authors

I. GENERAL SUBMISSION GUIDELINES

I.1. Manuscript Length

Articles should be between 6,000 and 8,000 words in length including all documents (abstract, keywords, tables, illustrations, acknowledgements, notes and references as well as text). Please indicate a word count.

I.2. Language:

I.2.1. Articles must be written in English. The authors themselves are reponsible for delivering their paper in good English.

I.2.2. Sexist language and idiomatic use of language should be avoided.

I.2.3. Spelling, capitalization and punctuation should be consistent within each article.

I.2.4. During the refereeing process authors should refer to their previous publications in the *third person*- not "as I stated in ..." but "as X noted ...".

I.2.5. Short and/or telegraph-style sections and/or paragraphs should be avoided.

I.3. Manuscript Format and Style

Contributors must submit **3 copies** of their manuscript (Word Perfect or MS Word).

I.3.1. Manuscripts should be left- justified only. Leave a 5 cm left margin, 2.5 cm elsewhere. Manuscript pages should be consecutively numbered (bottom centre). For the main body of the text double spacing should be used. The abstract, keywords, acknowledgement, notes, tables, references and appendices should be typed single-spaced. Indent the first line of each paragraph.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I.3.2. Each article should be supplied with (i) an abstract (no more than 150 words in length); and (ii) a maximum of 10 keywords or phrases, in English.

I.3.3. Elements of the article should be arranged as follows: title (written in bold type, capitalized and centred), author's full name (plus affiliation, mailing address, telephone and fax numbers, and e-mail address), abstract, keywords, text, acknowledgements, notes, references, and appendices if any. Please note that the author's name, affiliation, address, telephone/fax number and e-mail address should appear on only one copy to facilitate anonymous review.

I.3.4. Reports of empirical work should be structured conventionally: Introduction (including where appropriate a brief review of significant related work); Method (in sufficient detail to allow the study to be replicated); Results (only the most important results should be included); Discussion (avoiding a simple recapitulation of points already made); and Conclusions (related solely to the paper).

I.3.5. ***Main text***

Please, use the font Times New Roman (size 12).

I.3.5.a. Headings and subheadings

Up to four levels of headings are permissible. Please follow this style:

LEVEL 1: I. ROMAN NUMBERS, ALL CAPITALIZED, BOLD.

Level 2: I.1. Roman + Arabic numbers, bold.

Level 3: ***I.1.1. Roman + Arabic numbers, bold and italics.***

Level 4: *I.1.1.a. Roman + Arabic numbers + letter, italics.*

I.3.5.b. *Bibliographical citations in the text*

Bibliographical citations in the text must include the author's last name, year of publication and page references if applicable. Examples of correct styling for bibliographical citations are:

- Boughey (1997) / (Boughey, 1997) / Boughey (1998: 130) /(Boughey, 1998: 130).
- Connor and Kaplan (1987) / (Connor & Kaplan, 1987). Notice that when a reference is enclosed completely within parentheses the ampersand (&) is to be used instead of the word *and*. However, *and* is to be used outside the parenthesis.
- If more than one, citations should be listed in alphabetical order. Example: (Carson & Nelson, 1996; Cohen & Cavalcanti, 1990; Ferris, 1995; Hyland, 1990).

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I.3.5.c. *Quotations*

Quotations exceeding 100 words must be separated by a two-space gap from the main body of the text, indented from the left-hand margin and must be single-spaced. If not indicated in the main text, after the quotation indicate the author's last name, date of publication and page numbers in brackets. When shorter, double quotation marks ("...") must be used and be inserted in the main text. Special care must be taken to reproduce the originals exactly; any deliberate alterations must be indicated.

I.3.5.d. *Punctuation marks*

All punctuation marks (commas, semicolons, full stops) must follow any other text marks such as quotation marks (inverted commas) or (foot/end-)note numbering: example", example"; example". And also: example[7], example[7]; example[7].

If dashes are used instead of round brackets (in parentheses), they should appear immediately next to the beginning/ending word (with no space in-between), except at the end of a sentence preceding a full stop: this is just an example to —try to— illustrate this instruction. And: this is just another example to illustrate this instruction —or at least to try to.

I.3.6. *Notes*

They should follow the body of the text and be numbered consecutively throughout the text giving clear superscript numbers in the appropriate places. Notes should be avoided whenever this is reasonably possible.

I.3.7. *List of references*

This section should include the complete bibliographic information (*et al* must be avoided) of all the works cited and quoted in the text. Please follow the model given below paying special attention to punctuation, capitalization, spacing and indentation:

Journal Article:

Sánchez, A. (1992). Política de difusión del español. *International Journal of the Sociology of Language*, 95, 51-69.

Scheu-Lottgen, U.D. & Hernández-Campoy, J.M. (1998). An analysis of sociocultural miscommunication: English, Spanish and German. *International Journal of Intercultural Relations* (IJIR), vol. 22:4, 375-394.

Hilferty, J. Valenzuela-Manzanares, J. & Villarroya, O. (1998). Paradox Lost. *Cognitive Linguistics*, 9:2, 175-188.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Book:

Monroy, R. (1998). *Sistemas de transcripción fonética del inglés. Teoría y textos*. Granada: Grupo Editorial Universitario.

Calvo, C. & Weber, J.J. (1998). *The literature workbook*. London: Routledge.

Sánchez, A., Sarmiento, R., Cantos, P. & Simón, J. (1995). *CUMBRE. Corpus lingüístico del español contemporáneo: fundamentos, metodología y análisis*. Madrid: SGEL.

Edited book:

Barcelona, A. (Ed.) (2000). *Metaphor and metonymy at the crossroads: A cognitive perspective*. Berlin, New York: Mouton de Gruyter.

Pujante, A.L. & Gregor. K. (Eds.)(1996). *Teatro clásico en traducción: Texto, representación, recepción*. Murcia. Universidad de Murcia.

Book chapter:

Gregor, K. (1996). Unhappy Families: Ideology and English Domestic Life. In R. González (Ed.), *Culture and power: institutions*. Barcelona: Promociones y Publicaciones Universitarias, pp. 69-78.

Proceedings:

Conde-Silvestre, J.C. (2000). Variación estilística e isotopía textual en el poema anglosajón *The Ruin*. In *Selected papers in Language, Literature and Culture. Proceedings of the 17th International Conference of AEDEAN*, 321-25.

Martínez-Lorente, J. & Pujante, A.L. (2000). Mundos diatópicos y narrativa utópica: modelo social y conflicto humano en *We, Brave New World* y *1984*. In *Selected papers in language, literature and culture. Proceedings of the 17th International Conference of AEDEAN*, 383-87.

Doctoral Dissertation:

Walton, D. (1997). *Mail bondage. Sentencing Wilde between the sheets: An e*pest*emology of the epistolary (an architectonic rhapsody)*. Unpublished Doctoral Dissertation, University of Murcia, Spain.

Cantos-Gomez, P. (1995). *Lexical disambiguation based on standard dictionary definitions and corpus-extracted knowledge*. Unpublished Master Dissertation. University of Essex, Colchester, United Kingdom.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

Book Review:

Hernández-Campoy, J.M. (2000). Review of C. Silva-Corvalán (Ed.) (1995) *Spanish in four continents: Studies in language contact and bilingualism*. Washington, DC: Georgetown University Press. In *International Journal of Applied Linguistics* (InJAL), vol. 10.1, 141-145.

Technical Report:

Edwards, V. K., Trudgill, P. J. & Weltens, B. (1984). *The grammar of English dialects. A survey of research*. A Report to the ESRC Education and Human Development Committee. London: Economic and Social Research Council.

*Conference paper***:**

Gregor, K. (2000). *The torch and the marriage bed, or getting a grip on Cuchulain: The politics of Yeats's* On Baile's Strand. Paper given at *ESSE 5*, University of Helsinki, Finland, August.

Conde-Silvestre, J.C. & Hernández-Campoy, J.M. (1997). *The 'Acceptance' of the Chancery Standard in Late Medieval England: Sociolinguistic and Geolinguistic Tenets*. Paper given at the XXI International Conference of the Spanish Association of Anglo-American Studies (AEDEAN), University of Seville, Spain, December.

Manchón, R.M., Roca de Larios, J. & Murphy, L. (1997). *Lexical problems in L1 and L2 writing: Comparing beginner and intermediate foreign language learners.* Paper given at the XV Conference of the Spanish Association of Applied Linguistics (AESLA), Zaragoza, April.

I.3.8. ***Appendices***

Include the appendices starting on a separate page following the list of references. Each appendix should be labelled with numbers or letters and titled. If only one appendix is used, no identifying letter or number is required.

I.3.9. ***Tables, figures and graphs***

They should be titled and numbered consecutively throughout the article and appear as a unit after the reference section (or the appendices, if any). Prepare each table, figure or graph on its own page. At the point of the text where a table, etc. should appear, type an instruction such as "Insert Figure 1 here". In empirical studies the following information should be presented: (i) graphs and charts that explain the results; and (ii) complete source tables for statistical tests, when appropriate.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

I.3.10. ***Final page***

Please add a final page (not included in the word count) with the following information:

1) An *auxiliary short title* if the title of your paper exceeds 50 characters (including spaces). This will be used for running heads.
2) A list of figures, tables and graphs.

II. OTHER POINTS

II.1. Contributions should be submitted in typescript (3 copies) and PC or MAC diskette (WordPerfect or MS Word) to the specific Issue Editor(s) or the General Editor:

Pascual Cantos-Gómez
General Editor of *International Journal of English Studies*
Departamento de Filología Inglesa
Facultad de Letras
Universidad de Murcia
30071 Murcia
Spain

II.2. Deadline for submission of manuscripts: normally fixed by the Issue Editor and announced in the 'Call for Contributions'

II.3. Submission of a paper requires the assurance that the manuscript is an original work which has not been published previously and is not currently being considered for publication elsewhere.

II.4. Authors give the copyright to the publisher upon acceptance. Authors are also expected to take responsibility for obtaining permission to reproduce any illustrations, tables, etc. from other publishers.

II.5. Authors of accepted papers are responsible for proof-reading and must return proofs (via airmail, when appropriate) without delay. No modifications or additions will be accepted in the galley proofs except when completing references (i.e. for mentions such as: *in press*) or providing full reference to one's own previous work.

II.6. Copies: all author(s) will receive a free copy of the Journal.

II.7. Off-prints: The first-named author will receive 20 off-prints free of charge.

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

International Journal of English Studies

IJES

UNIVERSITY OF MURCIA www.um.es/engphil/ijes

Already Published Numbers

Number 1.1 (2001): ***Perspectives on Interlanguage Phonetics and Phonology***
Issue Editors: R. Monroy & F. Gutiérrez
Number 1.2 (2001): ***Writing in the L2 Classroom: Issues in Research and Pedagogy***
Issue Editor: Rosa M. Manchón
Number 2.1 (2002): ***New Trends in Computer Assisted Language Learning and Teaching***
Issue Editors: P.F. Pérez-Paredes & P. Cantos-Gómez
Number 2.2 (2002): ***Irish Studies Today***
Issue Editor: Keith Gregor
Number 3.1 (2003): ***Discourse Analysis Today***
Issue Editors: Dagmar Scheu & M.D. López Maestre
Number 3.2 (2003): ***Contrastive Cognitive Linguistics***
Issue Editors: J. Valenzuela & Ana Rojo
Number 4.1 (2004): ***Latest Developments in Language Teaching Methodology***
Issue Editors: Aquilino Sánchez & María Dueñas
Number 4.2 (2004): ***Advances in Optimality Theory***
Issue Editors: Paul Boersma & J.A. Cutillas-Espinosa
Number 5.1 (2005): ***Sociolinguistics and the History of English: Perspectives and Problems***
Issue Editors: J.C. Conde-Silvestre & J.M. Hernández-Campoy
Number 5.2 (2005): ***Editing Middle English in the 21st Century: Old Texts, New Approaches***
Issue Editor: Mª Nila Vázquez-González & J.C. Conde-Silvestre
Number 6.1 (2006): ***New Advances in Phraseological Research***
Issue Editor: Flor Mena
Number 6.2 (2006): ***Cognitive Phonology***
Issue Editor: J.A. Mompeán-González

Forthcoming Numbers

Number 7.1 (2007): ***Cognitive Linguistics: From Words to Discourse***
Issue Editors: Javier Valenzuela, Ana Rojo y Paula Cifuentes
Number 7.2 (2007): ***Research on L2 Vocabulary Acquisition and Learning***
Issue Editors: Aquilino Sánchez & Rosa María Manchón
Number 8.1 (2008): ***Software-aided Analysis of Language***
Issue Editors: P.F. Pérez-Paredes, M. Scott & P. Sánchez-Hernández
Number 8.2 (2008): ***A Survey of the Teaching and Learning of EFL in Spanish Speaking Countries***
Issue Editors: Pascual Cantos & Fernando Miño-Garcés
Number 9.1 (2009): ***Cultural Studies and Celebrity Culture: Contemporary Interventions***
Issue Editors: David Walton & John Storey
Number 9.2 (2009): ***Approaches to EFL Reading Comprehension: Research and Pedagogy***
Issue Editors: Piedad Fernández-Toledo & Françoise Salager-Meyer

© Servicio de Publicaciones. Universidad de Murcia. All rights reserved.

INTERNATIONAL JOURNAL OF APPLIED LINGUISTICS

Edited by
LEIV EGIL BREIVIK

The *International Journal of Applied Linguistics* seeks to cover all areas of applied linguistics and encourages the development of new fields of applied language study. The journal publishes original articles and reviews of current books (including books written in languages other than English), as well as notes and comments on points arising out of recently-published articles.

WWW.BLACKWELLPUBLISHING.COM/JOURNALS/INJAL

BLACKWELL PUBLISHING E-MAIL UPDATES

Advance notification of the latest articles in ***International Journal of Applied Linguistics*** e-mailed directly to your desktop. Join our free e-mail alerting service, and we'll send you journal tables of contents (with links to abstracts) and news of the latest books in your field.

SIGNING UP IS EASY. Simply visit: www.blackwellpublishing.com/ealerts

- Choose which discipline interests you, and we'll send you a message every two weeks
- OR select exactly which books and journals you'd like to hear about, and when you'd like to receive your messages.

E-mail alerts are also available for some journals through Blackwell Synergy.
For further details, visit: www.blackwell-synergy.com

Blackwell Publishing

Journal Customer Services, Blackwell Publishing,
PO Box 1354, 9600 Garsington Road, Oxford, OX4 2XG, UK. Tel: +44 (0) 1865 778315
customerservices@oxon.blackwellpublishing.com

visit our website for contents listings, abstracts, samples, and to subscribe
blackwellpublishing.com

JOURNAL OF SOCIOLINGUISTICS

Edited by
ALLAN BELL &
NIKOLAS COUPLAND

Now in four issues of 160 pages each per year, the *Journal of Sociolinguistics* has established itself as an international forum for multidisciplinary research on language and society.

The journal promotes sociolinguistics as a thoroughly linguistic and thoroughly social-scientific endeavour. The journal is concerned with language in all its dimensions, macro and micro, as formal features or abstract discourses, as situated talk or written text. Data in published articles represent a wide range of languages, regions and situations - from Alune to Xhosa, from Cameroun to Canada, from bulletin boards to dating ads.

WWW.BLACKWELLPUBLISHING.COM/JOURNALS/JOSL

BLACKWELL PUBLISHING E-MAIL UPDATES

Advance notification of the latest articles in ***Journal of Sociolinguistics*** e-mailed directly to your desktop. Join our free e-mail alerting service, and we'll send you journal tables of contents (with links to abstracts) and news of the latest books in your field.

SIGNING UP IS EASY. Simply visit: www.blackwellpublishing.com/ealerts

- Choose which discipline interests you, and we'll send you a message every two weeks
- OR select exactly which books and journals you'd like to hear about, and when you'd like to receive your messages.

E-mail alerts are also available for some journals through Blackwell Synergy.
For further details, visit: www.blackwell-synergy.com

Blackwell Publishing

Journal Customer Services, Blackwell Publishing,
PO Box 1354, 9600 Garsington Road, Oxford, OX4 2XG, UK. Tel: +44 (0) 1865 778315
customerservices@oxon.blackwellpublishing.com

visit our website for contents listings, abstracts, samples, and to subscribe
blackwellpublishing.com

Estudios de Sociolingüística
Linguas, sociedades e culturas

http://www.sociolinguistica.uvigo.es

Editors
Fernando Ramallo (Universidade de Vigo), framallo@uvigo.es
Xoán Paulo Rodríguez Yáñez (Universidade de Vigo), xoanp@uvigo.es

Reviews editor
Manuel Fernández Ferreiro (Universidade da Coruña), lxnanu@udc.es

About the journal

Estudios de Sociolingüística. Linguas, sociedades e culturas (EdS) will have two issues a year. EdS addresses a national and international specialized readership, and adopts a broad conception of the limits of sociolinguistics. This journal assumes as its own, the interdisciplinary character of sociolinguistics.

The Editors of EdS invite researchers to contribute papers and reviews to this journal. Certain issues of EdS will be monographical. EdS will accept proposals from contributors to prepare and edit collective monographical issues. EdS accepts contributions about traditional questions in this discipline and related fields such as: pragmatics, discourse analysis, conversational analysis, interactional linguistics, ethnography of communication, linguistic anthropology, ethnomethodology, language acquisition and socialization, etc. EdS recognises differences between schools and research orientations. Such orientations will constitute an equal part in this journal.

EdS will pay special attention to those theoretical and methodological contributions from sociology, social psychology or anthropology that contribute to give consistence to the theoretical and methodological principles of sociolinguistics, and the spread of conceptual evolution in the social sciences among sociolinguists.

Subscriptions

Please contact with:

SR. RICARDO SERRANO
MARCIAL PONS – DEPARTAMENTO DE REVISTAS
San Sotero, 6 — E-28037 – Madrid (Spain)
E-mail: nieves@marcialpons.es // serrano@marcialpons.es
Fax: 34 91 3272367 Phone: 34 91 3043303

2004 Subscription Rates (Two issues a year):

	Spain	*Rest of the World*
Institutions:	64 Euros ($66)	89 Euros ($92)
Individuals:	32 Euros ($33)	50 Euros ($52)

Postal address

Estudios de Sociolingüística
Facultade de Filoloxía e Tradución
Universidade de Vigo
Campus das Lagoas-Marcosende
E-36200 – Vigo (Galicia), Spain

Contents

IJES, Volume 6, Number 2, 2006

INTRODUCTION: COGNITIVE PHONOLOGY IN COGNITIVE LINGUISTICS vii
JOSÉ A. MOMPEÁN

ARTICLES

DAVID EDDINGTON
«Paradigm Uniformity and Analogy: The *Capitalistic* versus *Militaristic* Debate» 1

JOHN R. TAYLOR
«Where do Phonemes Come from? A View from the Bottom» 19

HELEN FRASER
«Phonological Concepts and Concept Formation: Metatheory, Theory and Application» 55

FUMIKI KUMASHIRO & TOSHIYUKI KUMASHIRO
«Interlexical Relations in English Stress» 77

GITTE KRISTIANSEN
«Towards a Usage-Based Cognitive Phonology» 107

JOSÉ A. MOMPEÁN
«The Phoneme as a Basic-Level Category: Experimental Evidence from English» 141

GEOFFREY S. NATHAN
‹‹Is the Phoneme Usage-Based? – Some Issues›› 173

BOOK REVIEWS

JOHN R. TAYLOR
‹‹Review of *Cognitive Phonology in Construction Grammar: Analytic Tools for Students of English* by Välimaa-Blum, Riitta. (2005). Berlin: Mouton de Gruyter›› 195

ABOUT THE AUTHORS 201

INSTRUCTIONS TO AUTHORS 205

ALREADY PUBLISHED AND FORTHCOMING NUMBERS 211

ISSN: 1578-7044